Manual
Archivista 2009/I

© 18th January 2009 by Archivista GmbH, CH-8118 Pfaffhausen
Web pages: www.archivista.ch

Contents

I Introduction **8**

1 Introduction **9**
- 1.1 Welcome to Archivista . . . 9
- 1.2 Notes on the manual . . . 9
- 1.3 Our address . . . 9
- 1.4 Previous versions . . . 9
- 1.5 Licensing . . . 12

2 First Steps **18**
- 2.1 Introduction . . . 18
- 2.2 The digital archive . . . 18
- 2.3 Database, server and client . . . 18
- 2.4 Tables, records and fields . . . 19
- 2.5 Archivista and working method . . . 19
- 2.6 Tips for archiving . . . 20
- 2.7 Archive, pages and documents . . . 20
- 2.8 The Archivista document . . . 20

3 Installation **22**
- 3.1 ArchivistaBox . . . 22
- 3.2 Virtual (Box) . . . 22
- 3.3 OpenSource (Box) . . . 22
- 3.4 OpenSource (Windows) . . . 22

II Tutorials **25**

4 Introduction **26**
- 4.1 How to use the online tutorials . . . 26
- 4.2 Installation ArchivistaBox . . . 26
- 4.3 Update an ArchivistaBox . . . 26
- 4.4 Accessing the manual . . . 26
- 4.5 Login WebClient . . . 26
- 4.6 Scanning and entering keywords . . . 26
- 4.7 Rotating pages . . . 26
- 4.8 Title search . . . 27
- 4.9 Full text search . . . 27
- 4.10 Login WebAdmin . . . 27
- 4.11 Adding users . . . 27
- 4.12 Adding/deleting fields . . . 27
- 4.13 Editing the input/search mask . . . 27
- 4.14 Activating SSH . . . 27
- 4.15 Activating VNC . . . 27
- 4.16 Enabling print server (CUPS) . . . 28
- 4.17 Password, Unlock & Restart OCR . . . 28
- 4.18 Activating HTTPS . . . 28

5 Tutorial RichClient **29**
- 5.1 Archivista in 90 Seconds . . . 29
- 5.2 Adding Documents . . . 29
- 5.3 Search . . . 29
- 5.4 Extended Functions . . . 30
- 5.5 Users & Fields . . . 31
- 5.6 Databases, fields and barcodes . . . 31

III ArchivistaBox **33**

6 Introduction **34**
- 6.1 Advantages . . . 34
- 6.2 DOLDER model . . . 34
- 6.3 RIGI and SAENITS model . . . 34
- 6.4 PILATUS model . . . 35

6.5 TITLIS model . . . 35
6.6 EIGER model . . . 36
6.7 MYTHEN model . . . 36
6.8 ROTHORN model . . . 36

7 Initial start-up 37
7.1 Introduction . . . 37
7.2 DOLDER, RIGI and SAENTIS . . . 37
7.3 PILATUS . . . 37
7.4 TITLIS and EIGER . . . 37
7.5 Switching it on . . . 38

8 System settings 42
8.1 Archivista WebClient (F10) . . . 42
8.2 Archivista WebAdmin (Alt+F9) . . . 42
8.3 Archivista WebConfig . . . 42
8.4 View manuals . . . 42
8.5 Archiving & OCR . . . 43
8.6 Archivista modules . . . 46
8.7 ArchivistaERP . . . 49
8.8 Backup . . . 49
8.9 Encryption . . . 51
8.10 Print server . . . 54
8.11 FTP Server . . . 56
8.12 Mail server . . . 59
8.13 Database . . . 60
8.14 System . . . 64
8.15 Remote Access . . . 73
8.16 Exit (Alt+F4) . . . 73
8.17 Function keys . . . 73

IV WebClient 75

9 User Manual WebClient 2009/I 76
9.1 Login . . . 76
9.2 Views . . . 76

10 Navigation 78
10.1 Main view . . . 78
10.2 Sorting . . . 78
10.3 Page view . . . 79

11 View, Search and Edit 80
11.1 View (Ctrl+F5) . . . 80
11.2 Search . . . 80
11.3 Edit mode . . . 82

12 Extended functions 86
12.1 Uploading a file . . . 86
12.2 Creating and scanning documents . . . 86
12.3 Editing and deleting pages . . . 87
12.4 Download . . . 87
12.5 Printing pages . . . 88
12.6 Barcode processing . . . 88

V WebAdmin 89

13 Login and Logout 90
13.1 Login . . . 90
13.2 Logout . . . 90

14 User 91
14.1 Administration (in general) . . . 91
14.2 User administration (external) . . . 94

15 Field definition 97
15.1 Field name . . . 97
15.2 Field type . . . 97
15.3 Length . . . 97
15.4 Position after . . . 97

16 Mask definition **98**
16.1 Mask . . . 98
16.2 Field name . . . 98
16.3 Field type . . . 98
16.4 Link with field . . . 101
16.5 Label . . . 101
16.6 Position . . . 101
16.7 Width . . . 101
16.8 User(s) allowed new entries . . . 101
16.9 User(s) allowed changes . . . 102

17 Archive administration **103**
17.1 The first options . . . 103
17.2 Further options . . . 105

18 Scan definitions **108**
18.1 Introduction . . . 108
18.2 General settings . . . 108
18.3 Settings . . . 108
18.4 Post editing . . . 110

19 Barcodes (Archivista Box) **113**
19.1 General . . . 113
19.2 Barcode technology . . . 113
19.3 Barcode entry . . . 114
19.4 Barcode recognition . . . 120
19.5 Barcode processing . . . 121

20 OCR definitions **123**
20.1 Name of OCR page definition . . . 123
20.2 Languages of definition . . . 123
20.3 Text quality of pages . . . 123
20.4 Further preparatory settings . . . 123
20.5 Options for table recognition . . . 123

21 SQL definitions **125**

22 Form Recognition (ArchivistaBox) **126**
22.1 Introduction . . . 126
22.2 Managing forms . . . 126
22.3 Editing objects of currently active definition . . . 127
22.4 Logo recognition . . . 130
22.5 Summary . . . 132

23 Exporting documents **133**
23.1 Introduction . . . 133
23.2 Exporting files in WebAdmin . . . 133
23.3 Exporting files in the WebClient . . . 133
23.4 Exporting files in the RichClient . . . 133

24 Mail archiving **134**
24.1 Introduction . . . 134
24.2 Individual definitions . . . 134
24.3 Editing a definition . . . 134
24.4 Possible problems during setup . . . 135

25 Database creation **136**
25.1 Option 'create' . . . 136
25.2 Option 'drop' . . . 136

VI WebConfig **137**

26 Administration with WebConfig **138**
26.1 Introduction . . . 138
26.2 Login . . . 138
26.3 Current Settings . . . 138
26.4 Setup scan button . . . 139
26.5 Backup . . . 139
26.6 Services . . . 140
26.7 Unlock documents . . . 140
26.8 Passwords ArchivistaBox . . . 140
26.9 Viewing log files . . . 141

26.10Text Recognition (OCR) . . . 141
26.11Turn off ArchivistaBox . . . 141

VII WebERP 142

27 ArchivistaERP 143
27.1 Introduction . . . 143
27.2 Activate ERP system . . . 143
27.3 First Step . . . 145
27.4 Create invoice . . . 147
27.5 Sales . . . 148
27.6 Purchases . . . 149
27.7 Items and Inventory . . . 149
27.8 Manufacturing . . . 149
27.9 Dimensions . . . 149
27.10Banking and General Ledger . . . 150
27.11Setup . . . 150

VIII RichClient 151

28 First steps 152
28.1 Before you begin... . . . 152
28.2 Starting up the program . . . 152
28.3 Creating a document . . . 152
28.4 Adding pages . . . 153
28.5 The main view of table 'Archive' . . . 154
28.6 Further tables of Archivista . . . 155

29 Menus, toolbar and shortcuts 156
29.1 Menus and functions . . . 156
29.2 Toolbar . . . 156
29.3 Shortcuts . . . 156
29.4 Tab key . . . 156
29.5 Resizing the detail view window . . . 156
29.6 Using the right mouse button . . . 157
29.7 Special fields . . . 157
29.8 Additional fields in table 'Archive' . . . 157

30 Menu Database 158
30.1 Printing (Ctrl+P) . . . 158
30.2 Working with databases . . . 158
30.3 Importing and exporting documents . . . 159
30.4 Hyperlinks (links between records) . . . 160
30.5 Parameters . . . 161
30.6 Page view (F9) . . . 161
30.7 OCR settings current page . . . 161
30.8 Quitting (Alt+F4) . . . 162

31 Menu 'Edit' 163
31.1 Working with the clipboard . . . 163
31.2 Working with abbreviations . . . 163
31.3 Navigating among the documents . . . 164
31.4 Create, copy, paste and delete documents . . . 164
31.5 Searching field content (Shift+F5) . . . 165
31.6 Replacing field content (Ctrl+F5) . . . 166
31.7 Creating documents & scanning pages (Shift+F9) . 167
31.8 Show pages . . . 167
31.9 Picture mode (Shift+Ctrl+F9) . . . 168
31.10Format (memo fields) . . . 168

32 Menu 'Search' and further menus 169
32.1 Documents (fields)...(F5) . . . 169
32.2 Pages (fulltext)... (Ctrl+F9) . . . 170
32.3 Sort (ascending) (Ctrl+F6) . . . 172
32.4 Sort (descending) (Ctrl+F7) . . . 172
32.5 All Records (F6) . . . 172
32.6 Working with selection criteria . . . 172
32.7 Menu 'Macros' . . . 173
32.8 Menu 'Table' . . . 173
32.9 Menu 'Help (F1)' . . . 174

33 Page View **175**
33.1 What is 'Page view'? . . . 175
33.2 Menu 'Archive' . . . 175
33.3 Menu 'Document' . . . 177
33.4 Menu 'Page' . . . 178
33.5 Menu 'Edit' . . . 183
33.6 Menu 'Viewing' . . . 186
33.7 Menu 'Macros' . . . 188

34 Further Tables **189**
34.1 Table 'Addresses' . . . 189
34.2 Table 'Notes' . . . 190
34.3 Table 'Literature' . . . 190

IX RichAdmin **192**

35 Parameters **193**
35.1 General program parameters . . . 193
35.2 User administration . . . 195
35.3 Directories (folders) . . . 201
35.4 Archiving & index... . . . 202
35.5 Internal image copies . . . 204
35.6 Field names . . . 204
35.7 Definitions for selection . . . 205
35.8 Edit macros... . . . 207

36 Define Fields and Masks **209**
36.1 Additional fields . . . 209
36.2 The form 'Field definitions and input masks' . . . 209
36.3 Working with fields . . . 210
36.4 First steps with objects . . . 211
36.5 Saving the form . . . 212
36.6 Field object: type . . . 212
36.7 Entry conditions . . . 218
36.8 Different input masks . . . 219

37 Define Print Reports **220**
37.1 Function . . . 220
37.2 Create new report . . . 220
37.3 Page setup . . . 220
37.4 Defining objects . . . 221

X Publishing **225**

38 PublishingEdition II **226**
38.1 A few tips on working with documents . . . 230

XI Technical Details **234**

39 External SQL calls (WebClient) **235**
39.1 External login . . . 235
39.2 Extended access . . . 235

40 External SQL calls (RichClient) **244**

41 Database structures **245**
41.1 The most important Archivista tables . . . 245
41.2 Archiving concept . . . 245
41.3 Clean login to Archivista databases . . . 249
41.4 File name of an external saved page . . . 249

42 ArchivistaBoxes **251**
42.1 BIOS settings . . . 251
42.2 Hard disk partitions . . . 252
42.3 Processing of new jobs . . . 252

43 RichClient **254**
43.1 Keywording functionality . . . 254
43.2 Efficient entering of data . . . 254
43.3 Runtime errors . . . 255

XII FAQs **256**

44 FAQs regarding ArchivistaBox **257**
44.1 Is there a demo version of Archivista? 257
44.2 ArchivistaBox does not start up properly 257
44.3 Shutdown of ArchivistaBox 258
44.4 Installation/Update from CD 258
44.5 Installation/update from USB-Stick 259
44.6 My CD does not boot 259
44.7 Installing ArchivistaBox on a Virtual Machine . . . 260
44.8 Saving your data 261
44.9 Data exchange and backup by means of USB stick 262
44.10LCD display with ArchivistaBox 264

45 Integration questions **265**
45.1 Is there a mail archiving function? 265
45.2 Can I archive Office files (incl. check out)? 265
45.3 How to scan from one ArchivistaBox to the other . 265
45.4 Scanning to ArchivistaBox with RichClient 266
45.5 Can I scan books? 268
45.6 Classifying documents (ERP interface) 269
45.7 Is there an API for the ArchivistaBox? 270
45.8 Import of images 270
45.9 Transferring the RichClient to the ArchivistaBox . . 272
45.10Importing data from other DMS software 272
45.11ArchivistaGPL (RichClient) 272

46 Solutions for common problems **275**
46.1 Scanning is slow 275
46.2 Scanner is not working 275
46.3 Characters on keyboard do not work 275
46.4 OCR is not (or no longer) functioning 275
46.5 ArchivistaBox is no longer working 276
46.6 Box is running but the monitor is black 276
46.7 Box is running but programs have vanished 276
46.8 A user has forgotten their password 276
46.9 'Document ... locked' message is displayed 276

Part I

Introduction

1 Introduction

1.1 Welcome to Archivista

We thank you sincerely for showing an interest in or even for buying Archivista 2009/I! After many years (how fast time passes) I take pleasure in welcoming you in a new preface. The Archivista products are at ten years and running.

During this time we have brought two major releases to the market. This does not look like much but this is exactly the point: continuity is where the emphasis is on. It started with refining our document structures internally for more than five years. This paid off well: not even the data structures for the archiving media did we have to change in the past decade.

Today we give you 30 years guarantee on our data structures. Again: for 30 years to come we will not touch the data structures of the archiving media (then, it will be over 40 years that we have the same structures and there is no reason why there should be an end to it...). We guarantee you the readability of our products and open data structures including source code (our new flagship product the ArchivistaBox is subject to the OpenSource license GPL) in order that you, your third party applications but also we ourselves may handle archived data trouble free.

This kind of constancy would not have been possible without You who contributed to the success of our products. We would like to take the opportunity to thank you here. In this sense I am looking forward to a continuously prosperous co-operation with you, today and in future.

Urs Pfister, Managing Director, Archivista GmbH

1.2 Notes on the manual

This manual is available on paper and as an online **help file.** The later can be started directly from within the program with the **function key 'F1'**.

➲ You can do a full text search on the online help file. Just enter your search string in the field at the top or the bottom of the dialog box and click on 'Search'.

Note on RichClient: Within the RichClient application there is also the readme file readme.txt. Where the versions Professional, Business and Enterprise differ a special note is made.

➲ Important notice or tip: whenever you see this symbol, you get relevant additional help.

1.3 Our address

You can reach us under the following address:

Archivista GmbH, Zürichstr. 80, CH-8118 Pfaffhausen
Tel: +41 (0)1 254 54 00; Fax: +41 (0)1 254 54 02,
E-mail webmaster@archivista.ch.

The latest version of Archivista is to be found at all times on www.archivista.ch.

1.4 Previous versions

1.4.1 Overview

From the mid-nineties when the Archivista product suite came into being a number of versions have been released. In chronological order:

1998: RichClient/RichAdmin: Endusers and Administrators work with the same program, the administrators merely possessing more access rights.

2002: WebClient: Endusers no longer need to retrieve and view documents via an installed program but can do it by means of a web browser. Advantage: greater flexibility, less implementation effort. However, administrators still make the settings by means of the RichClient/RichAdmin.

2005: Webadmin: Now endusers can also add documents and edit records via the web browser. Also administrators can make settings by means of a browser based interface.

2005: ArchivistaBox: Mini archiving server which is preconfigured with WebClient and WebAdmin. In addition, functions like backup, encryption, remote access, etc. are integrated in the ArchivistaBox.

2007: ArchivistaBox with AJAX (Web2): Ordinary web applications may be slower than RichClient applications because of the form concept. Thanks to new web technologies (AJAX in particular) it is possible to display content dynamically without having to reload forms. That is the reason why this technology is introduced with ArchivistaBox 2007/III. With version ArchivistaBox 2007/VIII also the RichClient is 100 percent OpenSource.

2008: ArchivistaBox as a BusinessBox: In addition to the web-based WebConfig application (web-based system settings), there is now an optional ArchivistaERP module. This now makes it possible to use the ArchivistaBox as a DMX Box as well as an ERP Box. Furthermore, the Linux port from Cuneiform provides good Open Source recognition which can also create searchable PDF files.

2009: More than 200 file formats can now be managed directly thanks to the integration of the Office formats. Documents can be sent to the archive and checked out again in their original versions. A more powerful mail archiving function is now also provided.

1.4.2 Versions in detail

1.4.2.1 ArchivistaBox 2009/I

The ArchivistaBox 2009/I can be used to send all Office documents (OpenOffice and Microsoft Office) to the archive (a function previously reserved for the old RichClient). Documents archived in this way can then be checked out again (in their original versions!) to undergo further processing. All formats (over 200 file types) can be processed without the need for external plug-ins. A more powerful mail archiving function is also provided which can be used to archive entire IMAP folders automatically. The system menu can now also be used to update the ArchivistaBox via the Internet. Split archive tables now make it possible to create archives of many terabytes in size.

1.4.2.2 ArchivistaBox 2008/XI

An optional LCD panel can now be ordered which can be used to trigger scanning via a small display. All settings can also be called from this display. In addition, the ArchivistaBox 2008/XI can now be run on netbooks (tested with the EeePC series). A new mobile scanner and the non-virtualised variant of the Dolder ArchivistaBox complement this perfectly.

1.4.2.3 ArchivistaBox 2008/IX

New Eiger ArchivistaBoxes are being introduced with the ArchivistaBox 2008/IX. On average, these are two to four times more powerful than the old Eiger Boxes. Good Open Source text recognition (Linux port from Cuneiform) is now also provided. In addition to text recognition, this can be used to create searchable PDF files in around 20 languages.

1.4.2.4 ArchivistaBox 2008/VI

French version: The WebClient, WebAdmin and the system interface are available in French. A new administration application for the system settings is provided with Archivista WebConfig. A new file upload is also supplied. The new API interface for the WebClient will probably also interest programmers. The web-based ArchivistaERP application is also new.

1.4.2.5 ArchivistaBox 2007/VIII

Beside various minor improvements there is now a comfortable import feature for images coming from a digital camera or a USB stick, respectively. Another new feature is form recognition which - together with logo recognition - enables the user to process documents automatically. Furthermore, the entire source code is now subject to the GPL license version 2. Lastly, ArchivistaBox can be attached directly to an LDAP server for central user management.

1.4.2.6 ArchivistaBox 2007/III

Totally new WebClient (Ajax based), improved scan engine (parallel processing of scan jobs), faster OCR on ArchivistaBox. Merging of documents (RichClient). Archived folders can be automatically written to CD/DVDs. All sessions can be protocolled in detail (WebClient) and viewed by means of the application AccessLog.

1.4.2.7 ArchivistaBox 2006/XI

Black/white optimisation for badly printed sheets of paper, splitting of documents (even without barcodes) during scanning process, automated writing of DVD/CDRW with complete archives, remote access for easy maintenance, multi core processor support with up to 4 GB RAM, automated creation of archives on USB sticks.

1.4.2.8 ArchivistaBox 2006/III

First released ArchivistaBox edition: print and FTP server. Backups on USB stick, tape, Windows/Linux file systems and USB hard disks. Master/Slave configurations. PDF/OCR server and barcode recognition module.

With ArchivistaBox 2006/III a new versioning concept begins. The year takes the first position followed by the month (in Roman Numerals). The software of the ArchivistaBoxes are subject to the GPL, i.e. the source code is open.

1.4.2.9 Archivista 5.2

¿From **version 5.2** onwards there is not only the RichClient version of Archivista but also the embedded OpenSource solution ArchivistaBox RIGI, PILATUS, TITLIS and EIGER. Smaller innovations include a photo mode for photo archives and an editor for templates in the barcode print program. (1 Oct 2005.)

1.4.2.10 Archivista 5.1

Version 5.1 offers many new features apart from a web module and a PublishingEdition for self-supporting archives. (15 April 2004).

1.4.2.11 Archivista 5.0

Version 5.0 embeds MySQL as database server and is able to handle much larger archives than hitherto. (30 March 2003).

1.4.2.12 Archivista 4.1

Version 4.16c works with FineReader5 engine (incl. PDF export) and offers post-editing functions for scanned documents (8 May 2002).

Version 4.15 introduces a fourth archiving format (JPeg). The archiving format may be specified on document level (15 January 2001).

Version 4.14 contains an enhanced FineReader4 engine and the new Search33 engine to index even larger archives. In addition to that there are many further enhancements (16 June 2000).

Version 4.12: FineReader 4.0 technology is introduced for OCR and a comprehensive barcode processing solution is incorporated too (19 July 1999).

Version 4.11 introduces two additional file formats. Along with that an English version is issued for the first time (30 Dec 1998).

1.4.2.13 Archivista 4.0

Version 4.0i was completely refurbished. In contrast to earlier versions Archivista 4.0i is no longer available as shareware version. (26 May 1998).

1.4.2.14 Archivista 3.x

Version 3.x with new modules 'Archive', 'Adresses', 'Notes' and 'Literature'. Version 3.x runs under Windows 9x/NT (1 June 1997).

1.4.2.15 Archivista 2.x

Version 2.x runs under Windows 3.x and Windows9x/NT (30 August 1996).

1.4.2.16 Archivista 1.x

Version 1.x: Internal version (20 January 1996).

1.4.3 Nobody is perfect

Archivista was developed over a period of more than ten years. Each version is amply tested before it is released. Nevertheless: nobody is perfect! Archivista GmbH is grateful for any indications regarding bugs or inconsistencies in the program functions.

1.5 Licensing

1.5.1 Licensing overview

Before the licensing types for specific products are described we would like to comment on why there are different licensing types at this juncture.

Archivista Version 4.x/5.x came into being under a standard proprietary license (ClosedSource). This means that the source code is not available to clients.

Three or four years ago we started with a new software generation that uses Internet related technologies and we introduced OpenSource tools. In principle we could have sold these particular bits under the GPL license even then. However, it did not seem to make much sense because it was not possible to work independently with the modules in question. It must be pointed out, too, that our RichClient products are subject to contracts which partly preclude the possibility to sell these applications under an OpenSource license.

On the occasion of the LinuxTag 2005 fair the software of the ArchivistaBox, the WebClient and the WebAdmin application was released under the GPL license because it is possible to work totally independently with it. In the year 2007 the entire source code of the RichClient, too, was made subject to the GPL license. For the rest it must be distinguished whether you download the software from our or any other homepage in the framework of the GPL or

whether you decide to buy the software from us or from authorized third parties and be entitled to support in return. If you buy an ArchivistaBox, for example, we offer support services. Should you then prefer to change from this support license to the GPL you are certainly free to do so.

1.5.2 Copyright note

Archivista GmbH, Zürichstr. 80, CH-8118 Pfaffhausen, possesses all rights regarding its software Archivista 2009/I.

Trademarks: MySQL (MySQL AB), Postscript & Acrobat (Adobe), Microsoft Word, Excel, Access and Windows (Microsoft), Apple Macintosh (Apple), Zweckform (Zweckform Büro Produkte GmgH), Ghostscript (Artifex) and Ghostgum (Ghostgum), AXIS (Axis Communications) and Xerox (Xerox Corporation). ABBYY(TM) FineReader(TM) OCR © 1993-2006 by ABBYY Software House. All rights reserved. ABBYY, FineReader are trademarks of ABBYY Software House. Part of imaging code: Pegasus Software LLC, Tampa, FL, www.pegasustools.com. Barcode from Softek Software www.softeksoftware.co.uk. VMware is a registered trademark of VMware, Inc., USA.. www.vmware.com.

Archivista 2009/I is protected by Swiss copyright law and international treaties.

1.5.3 Support for the ArchivistaBox

You can, of course, run the ArchivistaBox under the GPL license. You should note, however, that no support is available for this (see also GPL license under 1.5.6).

If you purchase an ArchivistaBox, you also receive the 'Silver' support package free of charge. This includes 30 days of telephone and email support for implementing the ArchivistaBox. You can also obtain free support by using our support forums and you will also receive the latest versions of the ArchivistaBox for three years. You can also choose to pay for support which you can do at any time over the course of five years. The rights of use of ArchivistaBoxes subject to GPL licenses are not affected by this.

If you require further support, a guaranteed reaction time or contractually agreed updates, then you require a maintenance contract in addition to the ArchivistaBox. If a maintenance contract is required, it must be concluded when the ArchivistaBox is purchased or retrospectively over the full period of time.

➲ If you still use 4.x or 5.x, we recommend that you update to the ArchivistaBox as soon as possible. The reason for this is that updated versions for 4.x and 5.x will no longer be available.

1.5.4 Guarantee

Archivista GmbH guarantees that the application Archivista 2009/I will run in basic accordance with this documentation for 90 days after selling date. The total liability of Archivista GmbH consists of a refund of the originally paid price or of subsequent improvement of the application (the choice of which is Archivista GmbH's). Misusing the program or manipulating it in a wrong manner voids this guarantee. There is no further guarantee. In particular, any liability with regard to consequential damages is excluded (as far as is lawful). Swiss law applies and the place of jurisdiction is Zürich.

1.5.5 Support and maintenance licenses

Should the basic support or the basic guarantee not be sufficient for your purposes, Archivista GmbH offers a variety of support and maintenance contracts. By buying an ArchivistaBox you are in any case entitled to a support license.

You cannot acquire a support license, however, if you use Archivista 2009/I under the GPL or test license. In such a case

we offer consulting. The conditions for consulting are indicated on our homepage.

1.5.6 The GNU General Public License

A considerable part of the Archivista software suite (ArchivistaBox, WebClient, WebAdmin, Publishing edition) is available under the GPL license. That is the reason why we reprint the original text of the GPL license.

Version 2, June 1991, © 1989, 1991 Free Software Foundation, Inc.,
51 Franklin St, Fifth Floor, Boston, MA 02110-1301, USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software—to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things.

To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

GNU General Public License

Terms and Conditions For Copying, Distribution and Modification

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The 'Program', below, refers to any such program or work, and a 'work based on the Program' means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term 'modification'.) Each licensee is addressed as 'you'.

   Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of

running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute verbatim copies of the Program's source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

   You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

   (a) You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

   (b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

   (c) If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

   These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

   Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

   In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

   (a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

   (b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to

be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

(c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

   If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

   It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to dis-

tribute software through any other system and a licensee cannot impose that choice.

This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

   Each version is given a distinguishing version number. If the Program specifies a version number of this License which applies to it and 'any later version', you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

No Warranty

11. Because the program is licensed free of charge, there is no warranty for the program, to the extent permitted by applicable law. Except when otherwise stated in writing the copyright holders and/or other parties provide the program 'as is' without warranty of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the quality and performance of the program is with you. Should the program prove defective, you assume the cost of all necessary servicing, repair or correction.

12. In no event unless required by applicable law or agreed to in writing will any copyright holder, or any other party who may modify and/or redistribute the program as permitted above, be liable to you for damages, including any general, special, incidental or consequential damages arising out of the use or inability to use the program (including but not limited to loss of data or data being rendered inaccurate or losses sustained by you or third parties or a failure of the program to operate with any other programs), even if such holder or other party has been advised of the possibility of such damages.

2 First Steps

2.1 Introduction

The following introduction provides you with some useful tips for getting started with an ArchivistaBox. If you want to start using the ArchivistaBox immediately, you can find the necessary information under 3.

2.2 The digital archive

From the time when information was first available in writing or in the form of pictures the problem arose that the knowledge thus collected would not be available when it was needed. This is where archiving sets in.

In contrast to a conventional filing system where every single page must be filed into folders or cabinets to be extracted later in a time-consuming search, Archivista helps you to do this work in a fraction of the time and in fully digitized manner.

This **time gain** is the result of well thought out processes of entering the material and searching for it effectively. Integrated **full text retrieval,** automated **barcode processing,** individual **subject allocation** or hyperlinks: all these functions are available to you in order that you may display any wished for document within seconds on your screen.

Digital archiving with Archivista means further that you can store both **paper documents** and documents that exist already in the computer as **electronic files.** Archivista is geared towards making information **available for** years – even **decades** – exactly in the same form as it existed in the original.

To effect this, all documents are stored in a picture format (before being compressed). Such picture formats (bitmaps) are very simple and therefore especially long living. Archivista works with **open and freely available standard bitmap formats** – a guarantee that your original documents will be correctly reproduced for decades to come.

It goes without saying that we could file all our documents in a tree directory on our hard disk. With a small archive this would work well. But the more documents we have in our archive the more difficult the administration of our data would be. This is the reason why Archivista 2009/I works with the high-performance database MySQL.

2.3 Database, server and client

At this point we would like to clarify a few terms whose understanding makes working with Archivista easier.

➲ If these terms are clear to you, please go directly to chapter 2.5.

Database: Organizational unit whose purpose it is to collect and retrieve information on the basis of particular criteria. A database could also be likened to a strong room where data are stored in a secure fashion.

Server and clients: In order to store our pieces of information safely we choose a safe location (server). This location may only be accessed by customers (clients) who have identified themselves as being rightful customers. To put it in a more abstract manner: the server offers clients services; in our case the service is a database.

Information exchange takes place in small quantities. If the customer (client) asks for the whole content of the strong room, he

gets the number of all available objects and the content of the first ten objects, but not the 990 other objects. If he wants to see more objects he must ask for objects 11-20. This helps to keep the strong room (server) uncluttered and make the exchange of data more efficient.

2.4 Tables, records and fields

If you can remember the following sentence you know almost all there is to know: a database contains tables, these contain records which themselves contain fields and in the fields a certain information is stored (e.g. a number).

Tables: The function of tables consists of pulling together the same type of information and making it visible (e.g. 'documents' in table 'Archive'). In Archivista there are other tables, e.g. table 'Addresses'.

Record: Each record contains a specified (identical) number of fields. Each row of the list represents one record and the top row contains the field names.

Fields: Each column in the list corresponds to a field. The field is the smallest organization unit and houses one piece of information (e.g. the date of the document).

2.5 Archivista and working method

Now that you are ready to take the plunge into the world of electronic archiving let us make a few remarks regarding Archivista and working method. Probably you have worked with folders hitherto, i.e. each document was filed in a folder. In the course of time you developed a folder structure. You might have had a separate folder for orders, invoices and reminders, for example. Alternatively, it may have been your clients which were the organizational unit of your filing system.

This **thinking in terms of folders is very inefficient from the point of view of digital archiving.** It simply does not make any sense to haul folders if all you are interested in is one document. Looking for a document may be a very time consuming task because neighboring documents not only guide your search but hamper it too. An index helps structuring your material but if you are forced to extend it, an often complicated reorganization follows suit.

That is the reason why we of Archivista GmbH have consciously opted against such time consuming folder structures. You will not find any imitated safes, cabinets, folders and registers in Archivista because their disadvantages in the realm of digital archiving exceed their advantages by far. Give your documents the information according to which you will search for them later. An invoice is named invoice and by adding a client name or client number and a date your document is uniquely defined. That is all.

When you look for your documents later on you do not have to dig in your archive for the appropriate folder but you simply tell Archivista that you want to see all invoices addressed to a specific client and that you want to see them sorted according to date. Within seconds you get a list of documents which correspond to these criteria. This is much faster than working your way through filing cabinets, folders and masses of documents in the same subfolder.

Do not only throw your old, dusty archive overboard but the working method that goes with it, too. Move on to the new advantages which digital archiving offers you. Enjoy!

2.6 Tips for archiving

What number of documents do you have? What sort of documents are they? The higher the number the more sense does a powerful scanner (including automatic document feeder) make. Apart from the velocity of the scanner the quality of the document feeder is highly important. Does it handle your specific (thin or thick) sheets without any problems? We gladly come up with suggestions for scanners which fulfill your particular needs. Please send an e-mail to webmaster@archivista.ch.

If your filing system relies on simple categories you must decide whether manual **subject allocation** makes sense or whether barcodes are a good idea. Manual subject allocation is suitable for up to 100 pages per day. If you have more paper to scan and archive you should think about barcode processing because barcode processing enables you to allocate subjects to your documents automatically. Thus all your documents are uniquely identified and perfectly categorized. Not a single one of them slips a later search because other than manual subject allocation which may result in faulty keywording now and then barcode processing is not prone to mistakes and it is faster, too.

Full text archive: If you scan many texts, the features text recognition and full text retrieval of Archivista are a boon. Do pay attention to choosing the right sort of language set. Umlauts are no problem then and thanks to the extensive dictionaries you may achieve excellent results.

Integration of documents already existing in digital form: Windows documents which already exist in electronic form may be incorporated into your archive by using DirectPrint2 (COLD). We recommend that you make yourself familiar with the basic functionality of Archivista before you try a hand at DirectPrint2. Please read chapter 8.10.

The **backup concept** is and remains vital in archiving. The archived pages should be copied to a write-once storage medium (like a CDR). Also the database itself as well as pages that are not yet archived should be backuped. So much on archiving in general (see chapter 26.5).

2.7 Archive, pages and documents

After installing Archivista the first **archive is created automatically.** You can begin feeding it with your papers straight away. There are two terms to know which is helpful: **documents and pages.**

Each sheet of paper corresponds to a **page.** Documents already existing in your computer are tied into the archive as if they had been printed, i.e. they are sent to a virtual printer and added to the archive.

A **document contains several pages** which belong together. A document may consist of only a single page but each page is allocated to a document.

➲ Each Archivista archive is made up of an undetermined number of documents. Each document consists of one or several pages.

Strictly speaking there is one more administrative unit: the folders. However, Archivista deals with them, you do not need to keep track of the folders. Only when discussing external storage devices will we touch on the topic of folders again.

2.8 The Archivista document

Each record in table 'Archive' is handled as a **'Document'**. The **document number** (field 'Document') is allocated automatically. Each 'Document' contains 640 pages at the most and will be

assigned to a **folder**. The **folder number** (Field 'Folder') is generated automatically according to the parameters set in 'Parameters'. At the beginning and as long as you do not transfer any data onto media other than your hard disk you work with folder number '1'. Later the folder number is automatically increased by 1 from time to time.

The fields 'Document', 'Pages' and 'Folder' cannot be edited (exception: system operator privileges). Neither can the fields 'Pages' and 'Archived' be edited. These show whether pages have been added and whether they have been archived respectively.

3 Installation

3.1 ArchivistaBox

You have purchased an ArchivistaBox. It is sitting in front of you and you want to start running it. To do so, see 7 for more information.

3.2 Virtual (Box)

You have also purchased an ArchivistaBox but you want to run it virtually. To do so, see 44.7 for more information.

3.3 OpenSource (Box)

➲ Please note that when you buy an ArchivistaBox the solution has been pre-installed on the harddisk. There is no installation necessary.

Should the need for an installation arise nevertheless, you will receive the entire system (ArchivistaBox, WebClient and WebAdmin) as an integrated package on a live CD or a live USB stick. After starting up the medium in question you simply run the installation process. Cf. 44.4.

3.3.1 Hardware needed

To work with it you need the following components:

- computer (256 MByte Ram or more, SCSI or USB2.0)
- harddisk (10 GB or more, recommended ¿40 GB)
- SANE enabled USB-/SCSI scanner

3.3.2 Getting CD and installing ArchivistaBox

The installation is simple. Firstly, download from www.archivista.chthe ISO file below (ca. 600 MByte) and create a CDR by means of a CD burning software. Then, start the CD on the computer you want to use as your archive. The live CD version starts up automatically. With this version you can test all functions before you make an installation on the harddisk of your computer.

3.3.3 Starting ArchivistaBox as liveCD

You will be asked for keyboard language and network port. After a few moments you will find yourself in the WebClient (full screen mode). To login use the password 'archivista' ('archivista' is the default password).

3.3.4 Installation from liveCD

To install the system on the harddisk leave the WebClient mode: press F12 to leave the full screen mode, rightclick your mouse and go to menu 'System'.. Run the function Install current system to a hard-disk partition.

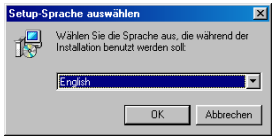
You have to determine the hard disk and confirm the installation. Depending on your computer this process will last 15 to 25 minutes. Subsequently, use F4 to shutdown the system ('archivista' must again be entered as password). Remove the live CD. Now, the system will boot from the hard disk.

3.4 OpenSource (Windows)

Since mid-2006, only one setup CD has been available, regardless of whether you want to work with the ArchivistaBox or the Rich-

Client. If you insert the CD during a Windows session, this usually initiates the setup program of the RichClient. However, if the CD is already in the drive when you start the computer and the drive is initiated by the CD, you are automatically taken to the LiveCD mode of the ArchivistaBox.

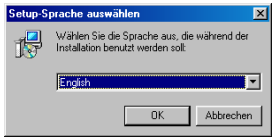
Quit all applications/programs and insert the Archivista CD into the CD-ROM drive. The installation program is started automatically. The first dialog box you see requires you to choose a language:



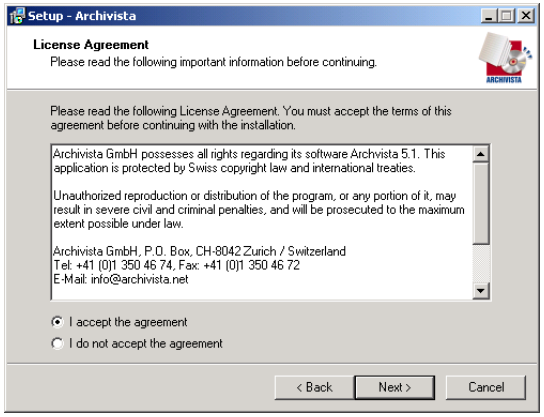
➲ The standard setup package contains all programs you need for the installation. This contrasts with version 5.0 where the individual programs had to be installed separately.

3.4.1 Step-by-step guidance

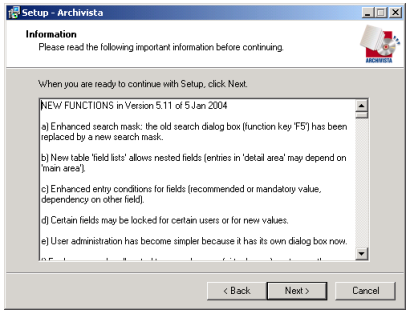
You need the current version of Archivista but at least version 5.2. During the setup procedure you will first see the following dialog box:



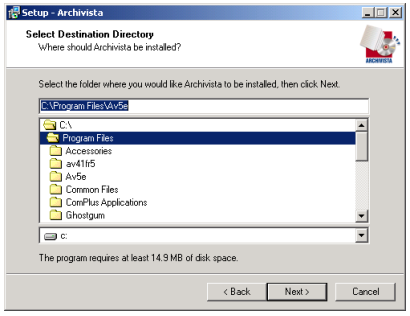
Please accept the license agreement for Archivista.


Setup - Archivista
License Agreement
Please read the following important information before continuing.
Please read the following License Agreement. You must accept the terms of this agreement before continuing with the installation.
Archivista GmbH possesses all rights regarding its software Archivista 5.1. This application is protected by Swiss copyright law and international treaties.
Unauthorized reproduction or distribution of the program, or any portion of it, may result in severe civil and criminal penalties, and will be prosecuted to the maximum extent possible under law.
Archivista GmbH, P.O. Box, CH-8042 Zurich / Switzerland
Tel: +41 (0)1 350 46 74, Fax: +41 (0)1 350 46 72
E-Mail: info@archivista.net
I accept the agreement
I do not accept the agreement
< Back
Next >
Cancel

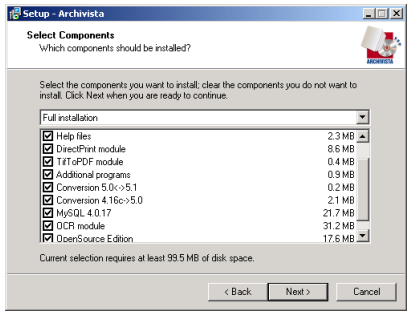
Please note the following tips and tricks:


Setup - Archivista
Information
Please read the following important information before continuing.
When you are ready to continue with Setup, click Next.
NEW FUNCTIONS in Version 5.11 of 5 Jan 2004
a) Enhanced search mask: the old search dialog box (function key 'F5') has been replaced by a new search mask.
b) New table 'field lists' allows nested fields (entries in 'detail area' may depend on 'main area').
c) Enhanced entry conditions for fields (recommended or mandatory value, dependency on other field).
d) Certain fields may be locked for certain users or for new values.
e) User administration has become simpler because it has its own dialog box now.
< Back
Next >
Cancel

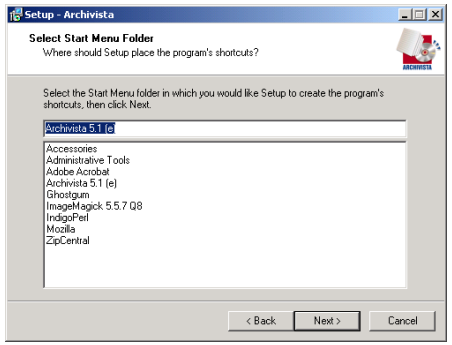
Now you must confirm the directory for the installation. Normally you simply click 'Continue.'



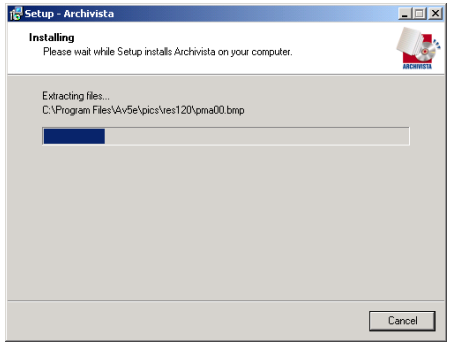
When choosing the components please make sure to tick the complete installation. Note too that during installation no MySQL Server must be running.


Setup - Archivista
Select Components
Which components should be installed?
Select the components you want to install; clear the components you do not want to install. Click Next when you are ready to continue.
Full installation
Help files 2.3 MB
DirectPrint module 8.6 MB
TiffToPDF module 0.4 MB
Additional programs 0.9 MB
Conversion 5.0<->5.1 0.2 MB
Conversion 4.16<>5.0 2.1 MB
MySQL 4.0.17 21.7 MB
OCR module 31.2 MB
OpenSource Edition 17.6 MB
Current selection requires at least 99.5 MB of disk space.
< Back
Next >
Cancel

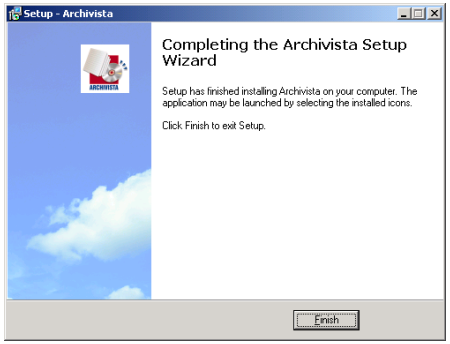
Now you confirm the folder for the start menu. Later you will find the Archivista program group under this entry.



Archivista is being installed; this process may take quite some time.



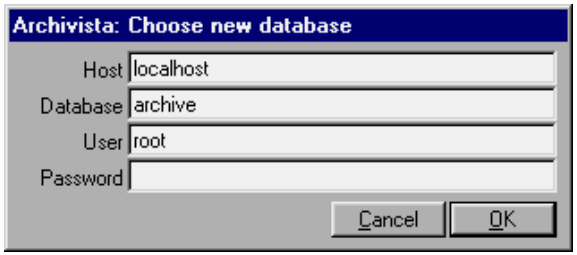
Finally, you see the following message:


Setup - Archivista
Completing the Archivista Setup Wizard
Setup has finished installing Archivista on your computer. The application may be launched by selecting the installed icons.
Click Finish to exit Setup.
Finish

Archivista 2009/I has been correctly installed.

3.4.2 Starting the program for the first time

After having installed Archivista 2009/I go to the 'Start' Menu in the lower left corner of your screen. In menu 'Programs' you will find 'Archivista 2009/I (e)' and within the range of programs appearing start **'Archivista 2009/I'.** The program is being started. After some time a dialog box appears which says: **'Archivista: Choose new database'.**



Now enter the connection data of the required ArchivistaBox database and confirm this by pressing the 'OK' button. After a few seconds, you are taken to the main Archivista screen.

➲: Please note that you have to create appropriate users in WebAdmin in the ArchivistaBox before you can use the RichClient to access the ArchivistaBox. See 14.1 for more information.

Part II

Tutorials

4 Introduction

4.1 How to use the online tutorials

To work efficiently with the Online Tutorial your screen should at least have a resolution of 1024x768 pixels. Each lesson consists of a GIF film and a short descriptive text (do not forget to scroll down).

Each film sequence lasts between 30 to 120 seconds. If you want to follow the lesson closely and go through the shown steps on your own in Archivista you will probably need about 10 minutes per lesson. At the end of each lesson you find a reference to the corresponding chapter in the manual. The combination of Online Tutorial and manual will give you maximum benefit.

Note: Since many clients have wished for it we slowed navigation time down for the purposes of the Online Tutorials. Thus you may distinguish the individual steps better. Real life navigation in Archivista is twice as fast as shown here. We hope to have effected for you an even more pleasurable plunge into the world of archiving.

Have fun with working through the Online Tutorial and with working with Archivista.

4.2 Installation ArchivistaBox

Tutorial on Installation of ArchivistaBox

In this tutorial you are shown how to install Archivista afresh on the ArchivistaBox whereby the existing configuration is overwritten.

Chapter(s) in manual: 44.4

4.3 Update an ArchivistaBox

Tutorial on Installation of ArchivistaBox part 2

This tutorial shows you how to update an ArchivistaBox and how to keep the configuration of the existing installation.

Chapter(s) in manual: 44.4

4.4 Accessing the manual

Tutorial on Accessing the manual

In the film you see how you can access the manual, jump directly to a particular chapter and move forward through the pages.

Chapter(s) in manual: 8.4

4.5 Login WebClient

Tutorial on Login to WebClient

Here you see how to log into the WebClient.

Chapter(s) in manual: 9.1

4.6 Scanning and entering keywords

Tutorial on Scanning and entering keywords

In this example a scanning process is triggered and subsequently a title keyword is added to the scanned document.

Chapter(s) in manual: 12.2

4.7 Rotating pages

Tutorial on Rotating pages

Pages in landscape mode that have erroneously been scanned portrait fashion can be rotated easily in Page view. In addition, wrongly scanned pages can be deleted.

Chapter(s) in manual: 12.3

4.8 Title search

Tutorial on Title search

Here it is shown how a search is carried out by entering a search term in the search mask.

Chapter(s) in manual: 11.2

4.9 Full text search

Tutorial on Full text search

The example here illustrates a full text search where a search term is entered, the result list is displayed and the user can jump easily from one occurrence of the search term to the next.

Chapter(s) in manual: 11.2

4.10 Login WebAdmin

Tutorial on Login to WebAdmin

This tutorial shows how a user can log into the administration tool 'WebAdmin'.

Chapter(s) in manual: 13

4.11 Adding users

Tutorial on Adding users

This film sequence teaches you how to add a new user. He/she is to gain access via web browser and should be required to enter a new password the first time he/she logs in.

Chapter(s) in manual: 14.1

4.12 Adding/deleting fields

Tutorial on Adding/deleting fields

Here you see how fields are added and deleted.

Chapter(s) in manual: 15

4.13 Editing the input/search mask

Tutorial on Editing the input/search mask

This example shows how a field is added to the input/search mask. At the beginning of the tutorial attention is drawn to the fact that an input/search mask is linked to user definitions.

Chapter(s) in manual: 16

4.14 Activating SSH

Tutorial on Activating SSH

Here SSH is activated for the purpose of establishing a secure connection to the console.

Chapter(s) in manual: 8.15.2

4.15 Activating VNC

Tutorial on Activating VNC. Tutorial on Accessing other PC by means of VNC

This film sequence shows how graphical remote access is enabled on the target PC and how you can subsequently access the target PC remotely.
Chapter(s) in manual: 8.15.1.1

4.16 Enabling print server (CUPS)

Tutorial on Enabling PDF printing
This film shows how the ArchivistaBox is 'opened up' to accept spool files that are sent to the archive from other PCs in the network. PDF printing can be turned off too.
Chapter(s) in manual: 8.10.1

4.17 Password, Unlock & Restart OCR

Tutorial on Changing password, unlocking documents and restarting OCR
Here you learn to do three things:

- how the root password is changed,
- how locked documents are unlocked,
- how OCR is restarted.

Chapter(s) in manual: 8.14

4.18 Activating HTTPS

Tutorial on Activating HTTPS
This example demonstrates how HTTPS (secure HTTP protocol) is turned on.
Chapter(s) in manual: 8.9.1

5 Tutorial RichClient

5.1 Archivista in 90 Seconds

Tutorial on Archivista in 90 seconds

In this example we do the following: first we create a document, then we go to 'page view', we scan two pages, we return to the main view, we allocate manually a subject to the document (with the help of the function 'abbreviations'), and we start the full text search.

Chapter(s) in manual: *28.3 Creating a document, 28.4 Adding pages, 8.5.3 Archiving documents*

5.2 Adding Documents

The following lessons teach you how to create documents and how to archive pages. The individual steps are the following.

5.2.1 Scanning pages

Tutorial on Scanning pages

In this example we scan two pages from within 'Page view', we delete superfluous information, we change the size of the viewed page and get back to the main view.

Chapter(s) in manual: *33 What is 'Page view'?, 33.4.2 Scanning with settings (F2)*

5.2.2 Keywording

Tutorial on Keywording

After adding pages you may allocate individual keywords to each document. To do this you use abbreviation functions. Within memo fields you may even format text bits!

Chapter(s) in manual: *31.2 Working with abbreviations, 31.10 Format (memo fields)*

5.2.3 Page view (part 1)

Tutorial on Page view (part 1)

Within page view you cannot only add pages but you can edit or delete them, too. You can also display them by twos ('Show two pages').

Chapter(s) in manual: *31 Menu 'Edit', 33.6 Menu 'Viewing', 33.6.14 Show two pages (F11)*

5.2.4 Working with JPEG

Tutorial on Working with JPEG

Working with JPEG ist very easy. Add a new document, choose the format and the pages you add will be stored in JPEG.

Chapter(s) in manual: *31.4.1 Create new document (Ctrl+N)*

5.3 Search

A powerful search functionality enables you to find archived pages quickly.

5.3.1 Searching and Replacing

Tutorial on Searching and Replacing

You may search your documents and pages not only with the help of the full text search feature but also with the help of the separately allocated subjects and other status information. You may use the menu 'Search' to make a first choice. The selected items may then be further searched or field values may be replaced.

Chapter(s) in manual: *32 Menu 'Search' and further menus, 32.1 Documents (fields) (F5), 31.6 Replacing field contents (Ctrl+F5)*

5.3.2 Full text retrieval

Tutorial on Full text retrieval

Thanks to the comfortable full text search function you may look through the entire body of text existing in your archive. Simply click the symbol of the magnifying glass or go to menu 'Search' and enter search words. By clicking the question mark in the search mask you learn the possible options for a detailed search. The hits may be displayed both in the main view and the 'Page view'.

Chapter(s) in manual: *32.2 Pages (fulltext) (Ctrl+F9)*

5.3.3 Editing page texts

Tutorial on Editing page texts

Since the quality of the text recognition depends heavily on the quality of the scanned pages mistakes cannot be avoided. Using a fabricated example we show you how simply you change the page text or make additions to it.

5.4 Extended Functions

The following four sections show you a small part of the manifold possibilites which you have when you work with Archivista.

5.4.1 Page view (part 2)

Tutorial on Page view (part 2)

Scanning occurs on the basis of one single form. After scanning there are functions which help you to edit your pages. You may rotate pages, delete areas of pages, delete entire pages to name only a few of the functions at your disposal.

Chapter(s) in manual: *33.5.1 Area (size), 31.4.4 Delete last page (Ctrl+L), 33.5.1.8 Divide page in two (Ctrl+T)*

5.4.2 Scanning brochures

Tutorial on Scanning brochures

Adding documents which are printed on both sides (front and back, e.g. brochures) used to require expensive duplex scanners. With the help of Archivista you may scan these pages with any scanner. We show you here how you add loose pages which are printed front and back. First we make a scan definition for sheets with the size A5. Afterwards we scan the front sides, then the back sides and finally we sort the pages in such a manner that they are in the right order again.

Chapter(s) in manual: *33.4.2 Scanning with settings (F2), 33.4.2.9 Definitions, 33.3.2 Sort*

5.4.3 Printing pages

Tutorial on Printing pages

A unique feature of Archivista are the screen copies. They are created when Archivista archives the pages and they are diminished to a size which fits the requirements of display on screen. This saves hard disk capacity and network resources because as a rule you need only the screen copies to work with. The full resolution copies of your pages are moved to CDRs or other external storage media from time to time and may be accessed that way. For printing purposes you may decide on the spot which quality you want just as you may decide about how many pages you want printed on one sheet. This saves time and paper.

Chapter(s) in manual: *33.2.1 Print (Ctrl+P)*

5.4.4 Print reports

Tutorial on Print reports

Apart from printing individual pages you may also print document lists and reports with the help of Archivista. You may freely define the reports you want just as you may send the printout to a file or to a faxdriver instead of a printer. In the case of labels you may choose the starting position.

Chapter(s) in manual: *37 Define Print Reports, 33.2.1 Print (Ctrl+P)*

5.5 Users & Fields

Archivista gives you the possibility to restrict user access in terms of individual users or groups. In addition, you learn in this chapter how to define your own fields, what field types there are and what is hiding behind the term 'field lists.'

5.5.1 Users and Groups

Tutorial on Users and Groups

You learn here how to create new users in Archivista. You can also create a group whose members all possess identical rights.

Chapter(s) in manual: *35.2 User administration*

5.5.2 Owner

Tutorial on Owner

The owner 'owns' the document. Neither the owner's nor his/her documents' existence need be visible to the other users. Compare the allocation of this kind of specific rights in the previous chapter.

Chapter(s) in manual: *35.2.1.8 New document with owner*

5.5.3 Adding new fields

Tutorial on Adding new fields

In this example you see how easily you can create new fields in your Archivista database. What is more, you learn how to tailor an input mask to your needs.

Chapter(s) in manual: *36 Defining Fields*

5.5.4 Type of Fields

Tutorial on Type of Fields

In this tutorial you learn how you can create fields which allow multiple entries. The program supports further types of fields (not shown here but described in the manual .

Chapter(s) in manual: *36.6 Field object: type*

5.5.5 Field lists

Tutorial on Field lists

In this example you are shown what field lists are for and how to change the order of the fields appearing in the Main View of your Archivista database.

Chapter(s) in manual: *36.6.2 Combo box, 36.6.3 Options for field objects*

5.5.6 Input conditions

Tutorial on Input conditions

It is possible on field level to reserve the right to make entries to select users. Apart from the privileged user no one may make changes to the content of the field.

Chapter(s) in manual: *36.7 Input conditions*

5.6 Databases, fields and barcodes

The following lessons are a step-by-step guide on databases and barcode printing.

Please note that the barcode processing requires the Barcode Option. The Barcode Option is only available with the versions Archivista Business and Archivista Enterprise.

5.6.1 Additional fields

Tutorial on Additional fields

To create additional fields run the function 'Define fields' in menu 'Database', submenu 'Parameters'. The mask helps you to create fields and adjust the detail view to your liking.

Chapter(s) in manual: *36.1 Additional fields*

5.6.2 BarcodePrint

Tutorial on BarcodePrint

The look of the BarcodePrint form is defined with the help of a configuration file. If you want to make an example of your own, you best use the values of this demo.

Chapter(s) in manual: *19.3.5 Define your own BarcodePrint forms*

5.6.3 Add barcodes

Tutorial on Add barcodes

With the help of the program 'BarcodePrint' you add barcodes easily and comfortably. The entry mask is displayed according to the previously created parameter file 'avbcode2.dat'. A probability check helps you when entering barcodes.

Chapter(s) in manual: *19.3.2 Working with Archivista-BarcodePrint*

Part III

ArchivistaBox

6 Introduction

The ArchivistaBox is a preconfigured server for everyday business use. It is focused towards document processing (DMS and archiving) as well as mapping business processes (ERP solution, new as of 2008). All user interfaces are web-based, the entire solution is delivered ready for operation in a single box and the sources comply with the GPL standard (i.e. the software is available as Open Source).

The following main models are supplied: DOLDER, RIGI, SÄNTIS, PILATUS, TITLIS and EIGER. There are also two sub-models that can be used as intelligent network scanning stations. These are called MYTHEN and ROTHORN.

6.1 Advantages

- No installation of any desktop application on an already busy server or on client workstations.
- Very intuitive user interface. Hardly any training necessary.
- Documents cannot only be searched but also edited via the web browser, i.e. keywords can be added, for example.
- The ArchivistaBox is independent of any operating systems already in use: it works together with Windows, Unix, OS/X etc.
- Starting to work with it is simple: plug and play!

6.2 DOLDER model

6.2.1 Hardware

The ArchivistaBox DOLDER is an industrial-suited fanless PC with the following content.

Geode LX800 Processor, 256MB RAM, fanless, 2xUSB, LAN, VGA, Compact flash harddisk with minimal 32GB (SSD), 12 Volt DC. Caused by the progress of hardware equipment components are subject to modifications.

6.3 RIGI and SAENITS model

6.3.1 Hardware

The ArchivistaBoxes RIGI and SAENTIS are industrial-suited fanless computers with the following content.

VIAC7 Processor, 256MB RAM, 1.2 GHz Processor, fanless. Interfaces: 2xUSB, LAN, VGA, 2.5 inch ATA HD with at least 32GB (SSD), 12 Volt DC. Caused by the progress of hardware equipment components are subject to modifications.

6.3.2 Scope of work ArchivistaBox Rigi

Single user version incl. OCR for a maximum of 20'000 documents per database. For a total of roughly 100'000 pages.

6.3.3 Scope of work ArchivistaBox Saenits

Multi user version incl. OCR for a maximum of 20'000 documents per database. For a total of roughly 100'000 pages.

6.4 PILATUS model

6.4.1 Hardware

The ArchivistaBox PILATUS is an industrial-suited PC with the following content.

Intel 865GV Core Logic Chipset, Intel Celeron 2GHz Processor, 512MB RAM, Automatic thermal fan control, Dimension: 14.8 x 25.4 x 7 cm, Interfaces: 1x1394, 3xUSB, LAN, serial, parallel, Line-in, Speaker, Headphone, Mic, VGA, 24x slim CD-Rom drive, 2.5 inch ATA HD with at least 160GB, 12 Volt DC. Caused by the progress of hardware equipment components are subject to modifications.

6.4.2 Software

Multi user version incl. OCR for three concurrent users. A maximum of 200'000 documents per database. For a total of roughly 1'000'000 pages. Users can optionally acquire further 'concurrent user' licenses or a barcode module license.

6.5 TITLIS model

6.5.1 Hardware

The ArchivistaBox TITLIS consists of two industrial-suited PCs that form a redundant system together. They have each the following content.

Intel 865GV Core Logic Chipset, Intel Celeron 2GHz Processor, 512MB RAM, Automatic thermal fan control, Dimension: 14.8 x 25.4 x 7 cm, Interfaces: 1x1394, 3xUSB, LAN, serial, parallel, Line-in, Speaker, Headphone, Mic, VGA, 24x slim CD-Rom drive, 2.5 inch ATA HD with at least 160GB, 12 Volt DC. Caused by the progress of hardware equipment components are subject to modifications.

6.5.2 Software

Multi user version incl. OCR for three concurrent users. For a maximum of 200'000 documents per database. For a total of roughly 2'000'000 pages. Users can optionally acquire further 'concurrent user' licenses or a barcode module license.

6.6 EIGER model

6.6.1 Hardware

The ArchivistaBox EIGER consists of two industrial-suited PCs that form a redundant system together and the backup is done by means of a tape drive. The PCs each have the following content.

Intel 865GV Core Logic Chipset, Intel Celeron 2GHz Processor, 1GB RAM, Automatic thermal fan control, Dimension: 14.8 x 25.4 x 7 cm, Interfaces: 1x1394, 3xUSB, LAN, serial, parallel, Line-in, Speaker, Headphone, Mic, VGA, 24x slim CD-Rom drive, 2.5 inch ATA HD with at least 320GB, 12 Volt DC. Caused by the progress of hardware equipment components are subject to modifications.

6.6.2 Software

Multi user version incl. OCR for three concurrent users. Unlimited number of documents per database. For a total of roughly 6'000'000 pages. Users can optionally acquire further 'concurrent user' licenses or a barcode module license.

6.7 MYTHEN model

The hardware corresponds to the RIGI hardware. The MYTHEN model serves exclusively as scan box, with which documents can be sent to the servers PILATUS or TITLIS. See 45.3

6.8 ROTHORN model

The hardware corresponds to the RIGI hardware. The ROTHORN model serves exclusively as scan box, with which documents can be sent to the server EIGER. Siehe 45.3

7 Initial start-up

7.1 Introduction

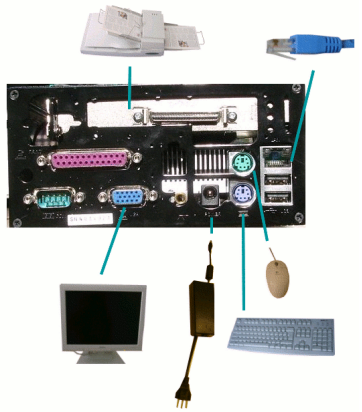
The ArchivistaBox is sitting in front of you and you want to get started. You will require a mouse, keyboard and a screen for the installation (with the exception of virtual ArchivistaBoxes). If you want to connect the ArchivistaBox with the outside world (i.e. a network), you will require a network cable and connection before you proceed.

7.2 DOLDER, RIGI and SAENTIS

The first step is to connect the ArchivistaBox with the input and output devices such as monitor, keyboard, mouse and scanner. Do not forget to plug in the power cables.

Attach peripherals and network cable.

7.3 PILATUS



➲ **Improtant: Use the USB interface at the front of the ArchivistaBox for scanning and backuping by means of the USB stick.**

7.4 TITLIS and EIGER

There are two boxes in front of you. You have to connect each box in the same way as for the PILATUS ArchivistaBox.

➲ Important: You can only connect a scanner to the master box. You must also ensure that you use the USB port on the front side of the master box.
See 8.13 for information about preparing the master/slave configuration.

7.5 Switching it on

Press the button 'Power' on the front of the ArchivistaBox. The box starts up and after about a minute you will get directly to the login screen of the Archivista web client.
➲ **Important:** If the screen flickers after you start the ArchivistaBox, then it has been unable to find a suitable resolution for the screen. In this case, disconnect the screen before switching on the ArchivistaBox. After two minutes, connect the screen. A resolution of 1024x768 is now in use. This procedure is also recommended if you connect the ArchivistaBox to a projector.

7.5.1 structure of the ArchivistaBox

There are four modes of operation in the ArchivistaBox.

- WebClient: Work as user (for example: add documents)
- WebAdmin: Make changes to an archive database (for example: edit fields)
- WebConfig: Make changes to the ArchivistaBox (for example: network address)
- System menu: Right-hand mouse button and 'Quit', right-hand mouse button again (only available in box)

If you want to work with ArchivistaERP, please see the information under 27.2.

The following describes how you access the various operating modes and what each of their purposes are.

7.5.2 Logging on to the WebClient

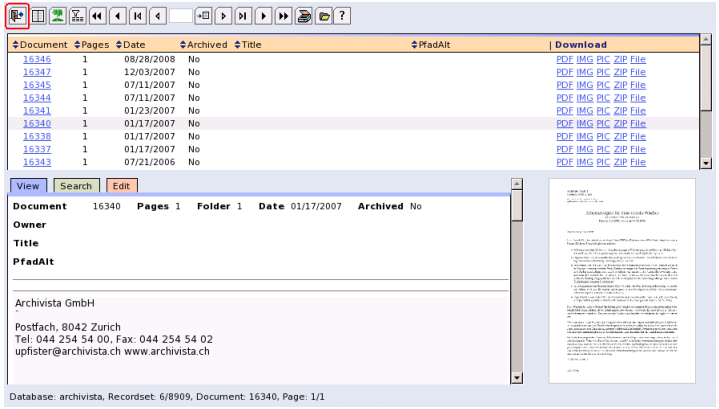
When you switch on the ArchivistaBox, the logon screen of the WebClient is displayed. The WebClient is designed for end users who work with archives. Documents can be entered and edited here.

To log in to the WebClient, enter the password and click on 'Login'.

➲ **Important:** The preset password is 'archivista' in lower case letters.

See the 9 chapter for information about the other functions of the WebClient.

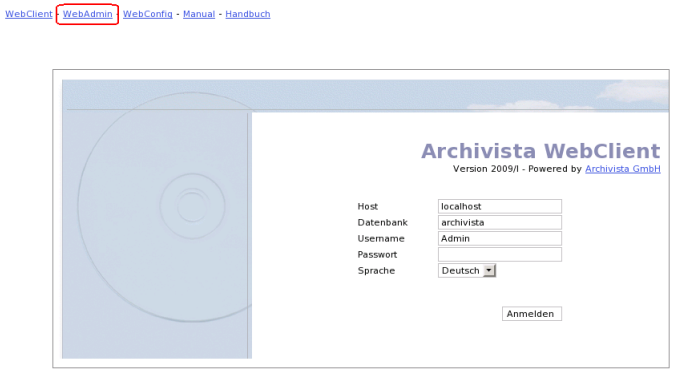
7.5.3 Exiting the WebClient


Document Pages Date Archived Title PfadAlt Download
View Search Edit
Document 16340 Pages 1 Folder 1 Date 01/17/2007 Archived No
Owner
Title
PfadAlt
Archivista GmbH
Postfach, 8042 Zurich
Tel: 044 254 54 00, Fax: 044 254 54 02
upfister@archivista.ch www.archivista.ch
Database: archivista, Recordset: 6/8909, Document: 16340, Page: 1/1

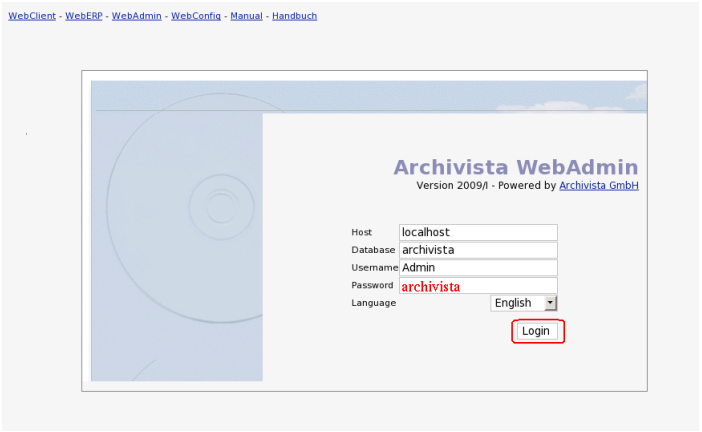
You can exit the WebClient by clicking on the first icon in the toolbar as shown below. You are then returned to the login screen.

7.5.4 Logging on to WebAdmin

When your ArchivistaBox is delivered, you will find that it includes the 'archivista' database which contains certain fields. Whenever you want to edit the structure of a database and/or you want to add or delete users, you have to call the WebAdmin application. Ensure that you are not logged in to the WebClient (see above). Then click on the corresponding 'WebAdmin' link on the upper left-hand side of the window.



You are then taken to the login window of the WebAdmin application:

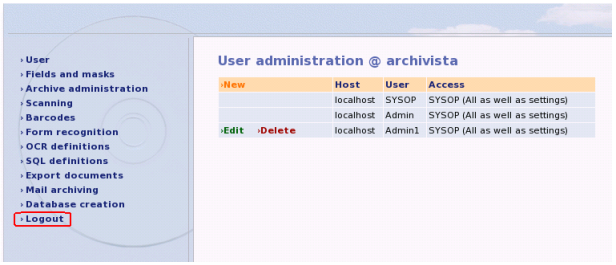


You also enter the password 'archivista' here. Then click on 'Login'.

The other functions of WebAdmin are outlined in the V section of this handbook.

7.5.5 Exiting WebAdmin

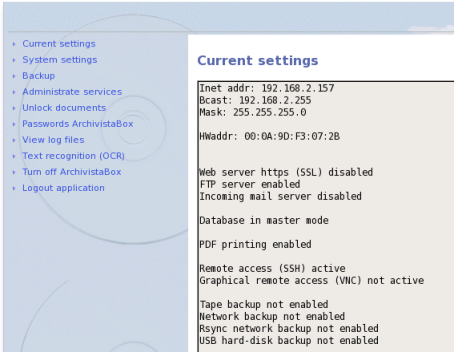
To exit the WebAdmin application, click on **'Logout'.** You are then returned to the login screen.



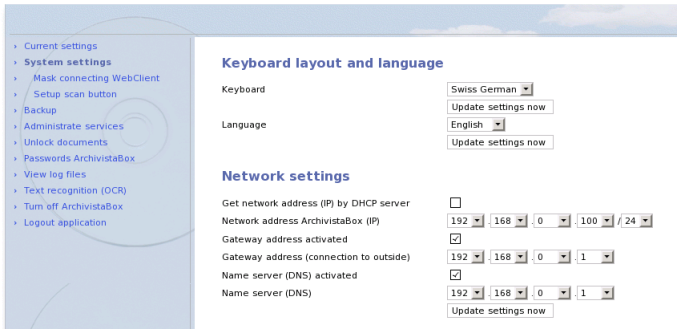
7.5.6 Logging in to WebConfig

In order to connect the ArchivistaBox to the outside world (i.e. a network), you have to call the 'WebConfig' application:

You also enter the password 'archivista' on the login screen here. Click on?Login'. You are then taken to the main screen:



Here, click on 'System settings' so that you can, for example, determine the network address:

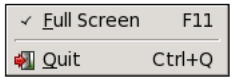


➲ **Important:** If the ArchivistaBox already has a network address, then it can be set from another computer if the WebConfig application is activated. If you do not want this, you can block the WebConfig application under 'System', 'Configure' and 'Configure WebConfig' in the system menu.

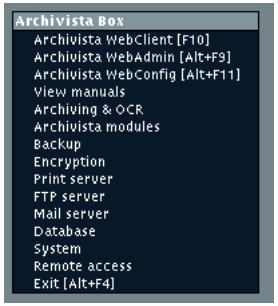
7.5.7 System menu

➲ **Important:** The system menu is only available in the ArchivistaBox itself. Thus, the following description only works if you are using an ArchivistaBox with a screen that is connected directly.

Once the ArchivistaBox has been started and is ready, the start screen for the WebClient application is displayed automatically. It may also be the case that an application has already been started. In order to now access the system menu, you have to close the application that is currently running. You do so by pressing the right-hand mouse button.



A grey screen is then displayed. Press the right-hand mouse button again to access the system menu.

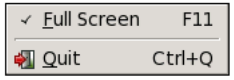


➲ **Difference between WebConfig and the system menu:** You can find almost all of the menu items of the system menu in the WebConfig application. Thus, you will rarely need to use the system menu. However, some menu items can only be accessed via the system menu (e.g. master/slave configuration). You also need the system menu if the WebConfig application has been deactivated.

8 System settings

You can make most settings in WebConfig - see 26. However, there are some settings (e.g. master/slave) that you can only make in the ArchivistaBox in the system menu. You must connect a screen, mouse and a keyboard to the ArchivistaBox to do this. Alternatively, you can use VNC to login directly in the ArchivistaBox. See 8.15.1 for more information.

If you are working in the ArchivistaBox itself, a web application (WebClient, WebAdmin or WebConfig) is usually active. In other words, you have to switch to the system level first. Press the right-hand mouse button first to do so. The following is then displayed:



Rightclick your mouse and click 'Quit'. An empty, grey screen appears. Richtclick the mouse and you see the following dialog window:

In the following we describe the commands one by one. Some functions are protected by a password. The default root password is 'archivista' written in lower case letters.

8.1 Archivista WebClient (F10)

Clicking this menu item takes you back to the Archivista WebClient.

8.2 Archivista WebAdmin (Alt+F9)

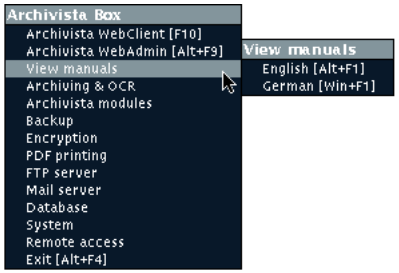
Clicking this menu item takes you back to the Archivista WebAdmin.

8.3 Archivista WebConfig

You can use Archivista WebConfig to manage the ArchivistaBox conveniently via a web application. Archivista WebConfig is described under 26.

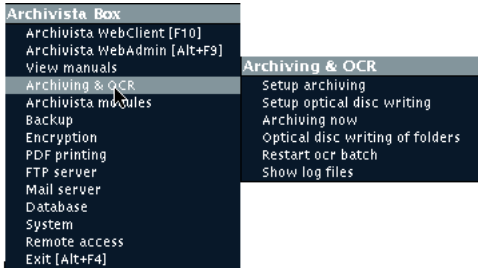
8.4 View manuals

By clicking 'View manuals' you get to the searchable PDF documents in English and German. Use the icon 'binoculars' to search for specific expressions.



Tutorial on Manuals (movie 2min) *4.4*

8.5 Archiving & OCR



The archiving process is part and parcel of every Archivista archive. This process is supposed to run at regular intervals. It keeps the archive tidy by 'rearranging' the newly added pages and generally does everything that needs to be done in order that documents and pages may be copied to external non-rewritable media.

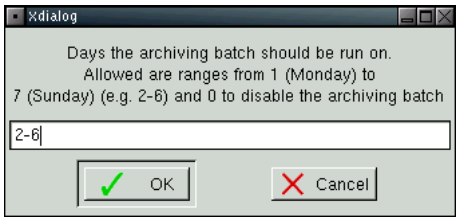
The options in connection with the archiving process are described in 17.2.6. This section is concerned with the actual triggering of the archiving process: when should it run and how often?

8.5.1 Setup archiving

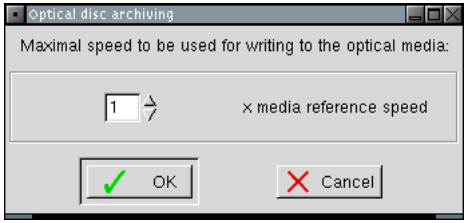
The root password must be entered first.

Enable archiving batch mode
Please enter the system password (root user) in order to enable the archiving batch mode.
Password:
Cancel
Run

Then you can specify on which days the archiving process is supposed to take place.


Xdialog
Days the archiving batch should be run on.
Allowed are ranges from 1 (Monday) to
7 (Sunday) (e.g. 2-6) and 0 to disable the archiving batch
2-6
OK
Cancel

Finally you can specify the time at which the archiving process should start.



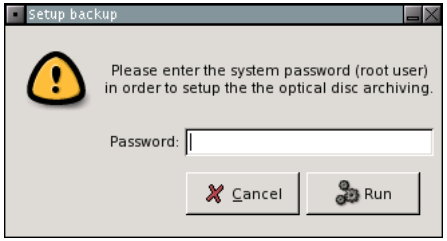
No confirmation message will be given. It is recommended to monitor the process by regularly consulting the log file 'Archiving & more' that is described in one of the following sections.

8.5.2 Setup optical disc writing

With this menu item you can configure attached CD/DVD writing devices so that later when the archiving process is run the folders in question are automatically transferred to CD/DVD.

➲ In case you would like to read an in-depth account of the Archivista archiving concept, you find the information in 41.

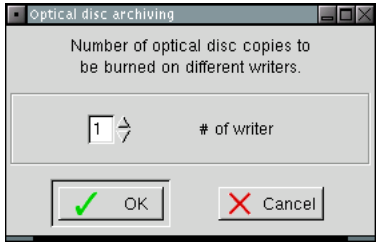
To arrange for automated creation of archiving folders you must get authorization first.


Setup backup
Please enter the system password (root user) in order to setup the the optical disc archiving.
Password:
Cancel
Run

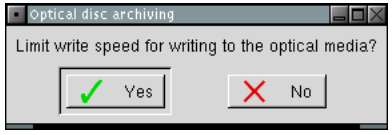
Now choose between 'CD' and 'DVD'.

➲ The decision for or against 'CD' or 'DVD' depends on several factors relating which would exceed the scope of this manual. In general one can say that the use of CDs (in particular gold layered CDRs) is still highly recommendable.

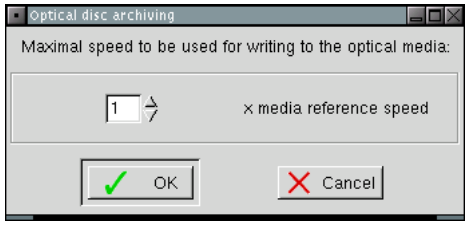
Now choose whether you have one or two CD/DVD writers at your disposal. We unreservedly recommend the simultaneous use of two writers as with this configuration you can always create two archiving media at the same time.



Determine whether you want to give an upper limit to the writing speed. Also this is highly recommended. For one, CD/DVD writers may not work properly at top speed, for another, the quality of the created media diminishes in proportion to the velocity with which they are written on – at least at the top end of the scale.



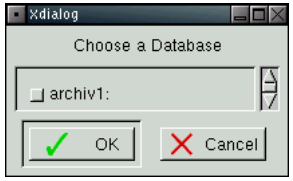
Set the desired writing speed. '1x' is admittedly too conservative for the majority of cases. When trying to find the optimum speed keep in mind the remarks above.



With this the configuration of the CD/DVD writers is finished. You can start the archiving process and wait until the first medium is written. It goes without saying that you must first put a CD/DVD medium into the CD/DVD device. And yes, depending on the number of documents the writing process may take longer. On average, a CD/DVD should be written after a couple of thousand pages.

8.5.3 Archiving now

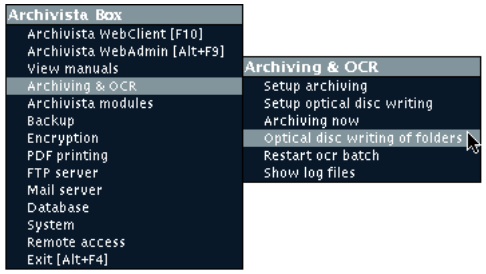
Alternatively, the archiving process can be triggered immediately. After having entered the root password you can pick the databases for which the archiving process is to be run. (The example below shows only one database.)



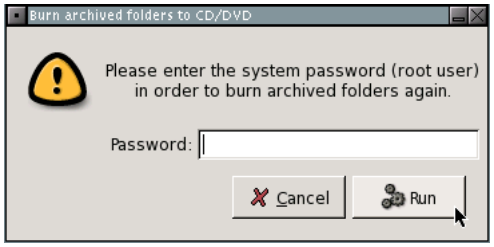
No confirmation message will be given. It is recommended to monitor the process by consulting the log file 'Archiving & more' that is described in one of the following sections.

8.5.4 Optical disc writing of folders

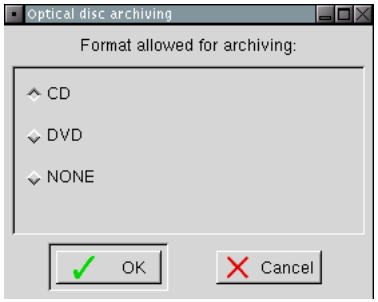
The purpose of the function **Optical disc writing of folders** is that the folders are not only saved in the database but on the box. This means that the folders are saved externally for long-term accessibility. The folders saved externally are written to a CD or DVD. To make a copy of one or several folders proceed in the following manner:



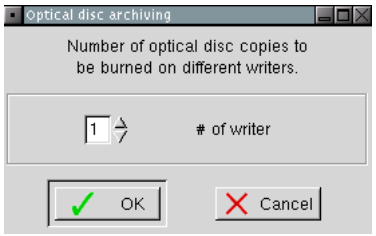
Choose function **Optical disc writing of folders** in menu item **Archiving&OCR**.


Burn archived folders to CD/DVD
Please enter the system password (root user) in order to burn archived folders again.
Password:
Cancel
Run

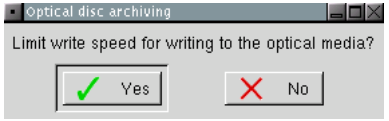
Enter the password.



Choose the medium on which you want to archive the data.

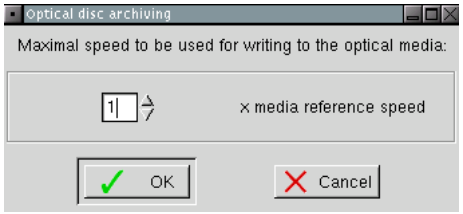


Enter the number of copies in the field. E.g. '1'.



Choose here if you want to restrict writing speed.

➲ This is to be recommended in any case as a writing speed that is too high may impair data quality or may even damage the medium.



Choose here the maximum speed for the writing procedure.

8.5.5 Restart OCR batch

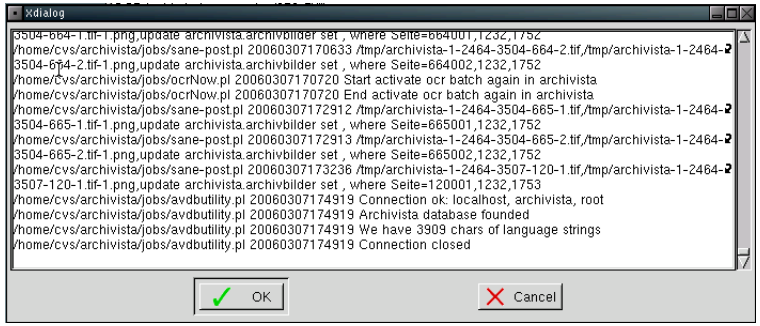
This function starts the OCR as batch program. I.e. the OCR is not running continually in the background but is started manually. No confirmation message will be given. The log files described in the next section serve to monitor the process.

➲ In order that the text recognition software may run smoothly the ArchivistaBox must be closed down once and started up again.

8.5.6 Show log file

8.5.6.1 Archiving & more

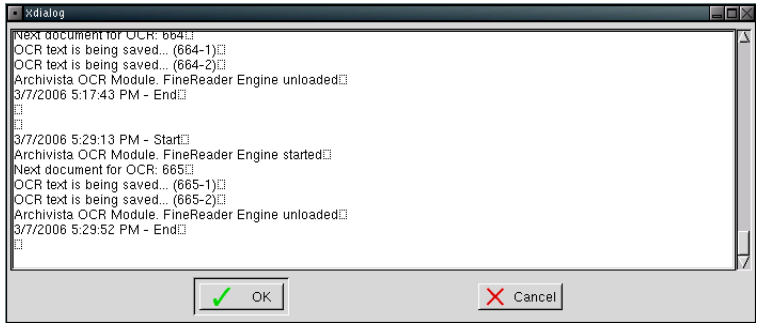
The log file gives information about whether the archiving process has been started and about the procedures already finished as well as about potential inconsistencies. There are two submenus, namely 'Archiving & more' and 'Text recognition'.


Xdialog
/home/cvs/archivista/jobs/sane-post.pl 20060307170633 /tmp/archivista-1-2464-3504-664-2.tif,/tmp/archivista-1-2464-2
3504-664-2.tif-1.png,update archivista.archivbilder set , where Seite=664002,1232,1752
/home/cvs/archivista/jobs/ocrNow.pl 20060307170720 Start activate ocr batch again in archivista
/home/cvs/archivista/jobs/ocrNow.pl 20060307170720 End activate ocr batch again in archivista
/home/cvs/archivista/jobs/sane-post.pl 20060307172912 /tmp/archivista-1-2464-3504-665-1.tif,/tmp/archivista-1-2464-2
3504-665-1.tif-1.png,update archivista.archivbilder set , where Seite=665001,1232,1752
/home/cvs/archivista/jobs/sane-post.pl 20060307172913 /tmp/archivista-1-2464-3504-665-2.tif,/tmp/archivista-1-2464-2
3504-665-2.tif-1.png,update archivista.archivbilder set , where Seite=665002,1232,1752
/home/cvs/archivista/jobs/sane-post.pl 20060307173236 /tmp/archivista-1-2464-3507-120-1.tif,/tmp/archivista-1-2464-2
3507-120-1.tif-1.png,update archivista.archivbilder set , where Seite=120001,1232,1753
/home/cvs/archivista/jobs/avdbutility.pl 20060307174919 Connection ok: localhost, archivista, root
/home/cvs/archivista/jobs/avdbutility.pl 20060307174919 Archivista database founded
/home/cvs/archivista/jobs/avdbutility.pl 20060307174919 We have 3909 chars of language strings
/home/cvs/archivista/jobs/avdbutility.pl 20060307174919 Connection closed
OK
Cancel

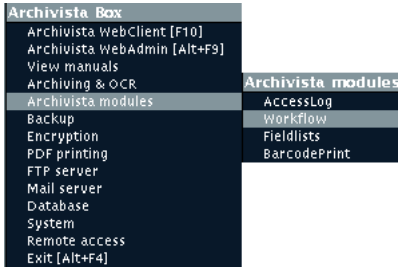
From this particular log file of 'Archiving & more' it can be seen that the OCR process was started and ended.

8.5.6.2 Text recognition

In the log file 'Text recognition' details of the text recognition process are listed.

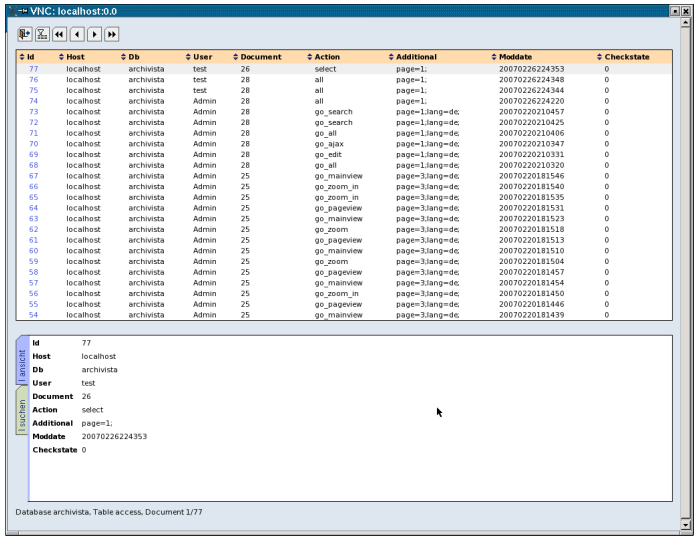

Xdialog
OCR text is being saved... (664-1)
OCR text is being saved... (664-2)
Archivista OCR Module. FineReader Engine unloaded
3/7/2006 5:17:43 PM - End
3/7/2006 5:29:13 PM - Start
Archivista OCR Module. FineReader Engine started
Next document for OCR: 665
OCR text is being saved... (665-1)
OCR text is being saved... (665-2)
Archivista OCR Module. FineReader Engine unloaded
3/7/2006 5:29:52 PM - End
OK
Cancel

8.6 Archivista modules



8.6.1 AccessLog

This module serves to create a logfile of all accesses to the Web-Client.



8.6.1.1 Turning it on

Please note that this option is normally not turned on. You must activate it as described in 17.2.9.

8.6.1.2 RichClient

The RichClient is currently not able to access archives for which activities are logged in detail. When trying to log in you will get an error message.

8.6.1.3 Slave computer

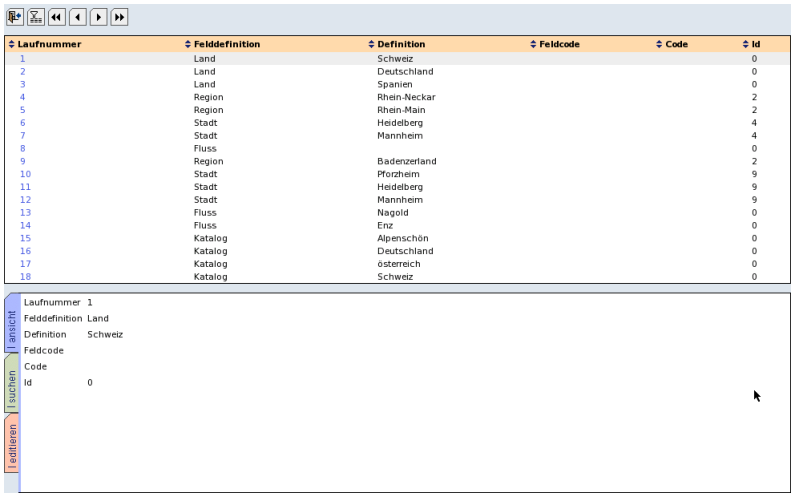
Please note that it is not possible to access an archive on the Slave Box by means of the WebClient (as all accesses can be protocolled only once). To nevertheless ascertain whether the ArchivistaBox in Slave mode does cleanly replicate the Master Box you will see, when logging into the Slave computer, an error message displaying the last modification that has taken place in the database.

➲ The following may serve as example: Try to log into the Slave. The time appears when the last change was made. Now go back to the Master and use the WebClient to navigate from one document to the next. This causes the 'AccessLog' table to be automatically updated in the background. Now go back to the Slave and use the Webclient to login. The generated time stamp must be different from the first because the activity registered by the 'AccessLog' table in turn influences the ArchivistaBox Slave.

8.6.2 Workflow

The module 'Workflow' is not active for all customers, i.e. it is not preconfigured on the OpenSource CD and ready to use. We are working on it...

8.6.3 Fieldlist (ArchivistaBox)



8.6.3.1 Introduction

The fieldlist represents a kind of central management of field entries. Go to fieldlist by means of menu `Archivista Modules`. Here all entries of almost all `fields` created in `WebAdmin` are listed. Fields of field type `Normal` do not appear in the fieldlist as they are not connected with any other field. In addition to a simple listing also the hierarchies between the individual fields are shown. Thus, the field list gives an overview over the relations between the existing fields.

➲ However, special care must be taken: editing the `fieldlist` requires considerable circumspection as the entire tree might be destroyed if one makes editing mistakes.

8.6.3.2 The navigation

There are six navigation symbols:

Quit program (Alt+q)

Select all records (Alt+a)

Previous records in selection (Alt+f)

Previous record (Alt+p)

Next record (Alt+n)

Next records in selection (Alt+l)

8.6.3.3 Representation of different field types

To give as simple an overview as possible the tree structure is visualized in a table. The field types are represented in different ways depending on their dependency status. Field type `Definition` and field type `Normal` are at the top of the field type hierarchy. As mentioned before field type `Normal` is not shown in the `fieldlist`. Hence, field type `Definition` is at the very top of the field types. All other field types are subordinated to this particular type and exist in relation to a `Definition`. There are six different field types. For more detailed information see Section 16.3.

The fieldlist is structured in six columns:

- Sequence number
- Field definition
- Definition
- Field code
- Code
- ID

The `Sequence number` mirrors the order of the different field type definitions in the fieldlist.

The column `Field definition` shows those fields that were allocated either the field type `Definition` or the field type `1:N` in menu mask definition of **WebAdmin**.

Example: The field `Country` is allocated the field type 'Definition'. Hence, `Country` will appear in column `Field definition`. The countries allocated to the field `Country` in the **WebClient** appear in column `Definition`. If we assume that the field `Country` has the sequence number '1' in the fieldlist, this sequence number appears also beside the field `Region` that can also be found in column `Field definition`; `Region` is of the field type `1:N` that appears in column `ID`. Explanation: the region was subordinated to the country by allocation of field type `1:N`. The entries made in field `Region` in the **WebClient** appear in column `Definition`.

Laufnummer	Felddefinition	Definition	Feldcode	Code	Id
1	Land	Schweiz			0
20	Region	Zürich			1

Fields to which the field types `Text` or `Number` have been allocated appear in column `Field code`. The number code to which a code was allocated in the **WebClient** appears in column `Code`. For more information on how to edit fields in the **WebClient** go to Chapter 11.3

Example: We are to develop a system whereby each company is assinged a number. For the purpose of illustration the name

of the sample company is **Petermann AG**. We first have to create the field `Company name` and then assign it to field type `Definition`. This field appears in the `Fieldlist` in column `Field definition`. In the **WebClient** we make the entry **Petermann AG** in field `Company name`. This entry can then be found in column `Definition`. To assign a company number to this company one creates the field `Company number` and allocates the field type `Number code` to it. The codes pertaining to the companies are assigned in the **WebClient**. Subsequently they will appear in column `Code`.

Laufnummer	Felddefinition	Definition	Feldcode	Code	Id
21	Firmenname	Petermann AG	Firmennummer	2	0

The idea behind the `Multi field` is to be able to use more than one keyword of the same order for one document. The field type `Multi` is also subordinated to the field type `Definition`. Hence, each `Multi field` must be linked to a `Definition field`. Compare Chapter 16.3.6.

In the case of a `Multi field` the field with the definition type `Definition` appears in column `Field definition` and the field with definition type `Multi` in column `Definition`.

Laufnummer	Felddefinition	Definition	Feldcode	Code	Id
16	Katalog	Deutschland			0
17	Katalog	österreich			0
18	Katalog	Schweiz			0

In this example 'catalog' was given the field type `Definition`. For the multi fields `catalog1` and `catalog2` Germany, Switzerland and Austria were added as entries.

8.6.3.4 The mask

In the lower part of the screen there is a mask with three different tabs:

- View
- Search
- Edit

`View` shows the information that is contained in the fieldlist. The `Search` mode enables the search for strings or a group of strings on the basis of individual characteristics. For example, if one wanted to know which records are linked to the record with sequence number '3' one enters '3' in `ID` in the search mask and subsequently the system returns all records subordinated to record '3' by means of field type `1:N`.

Behind tab `Edit` there is a window in the upper left corner that gives us the options to add new strings or to change existing strings. In addition, existing information can be changed.

8.6.4 BarcodePrint

The module 'BarcodePrint' is not active for all customers, i.e. it is not preconfigured on the OpenSource CD and ready to use.

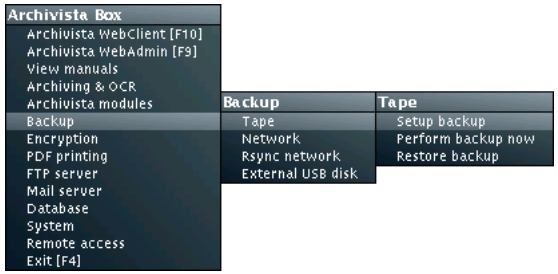
8.7 ArchivistaERP

These options are described in detail under 27.

8.8 Backup

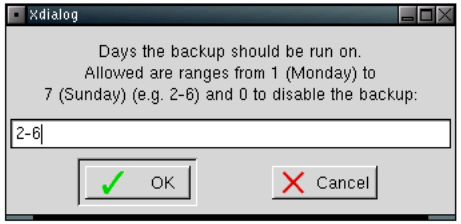
The purpose of this menu item is to make regular backup copies of your data to a different drive. There are four alternatives, which are described below.

8.8.1 Tape

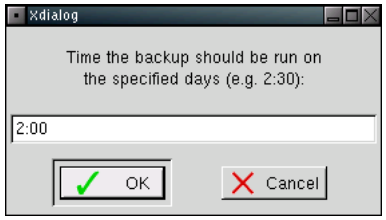


8.8.1.1 Setup backup

First, you will be asked for the system password. Subsequently, the following dialog box will appear:


Xdialog
Days the backup should be run on.
Allowed are ranges from 1 (Monday) to
7 (Sunday) (e.g. 2-6) and 0 to disable the backup:
2-6
OK
Cancel

Enter the days when you want a backup to take place.



Enter the time of day when the backup should take place. The default value '2:00' means 2 o'clock in the morning.

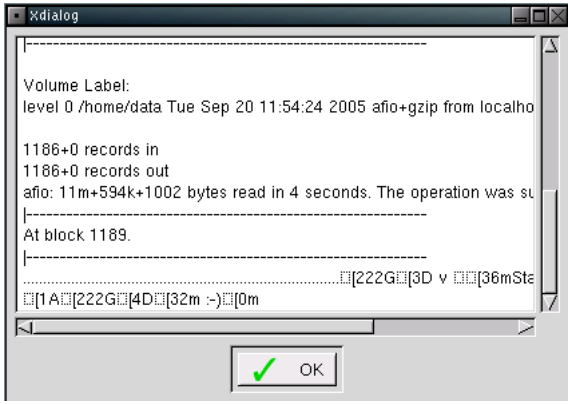
➲ There is no further confirmation. After the time of day has been entered the backup is properly setup. Each time the backup process has successfully ended (usually each morning) you will get a message to that effect.

8.8.1.2 Perform backup now

The menu item 'Perform backup now' kicks off an immediate backuping procedure. Click 'Perform backup now', enter the system password and the copying process starts under the condition that a tape drive is attached to the ArchivistaBox.

8.8.1.3 Restore backup

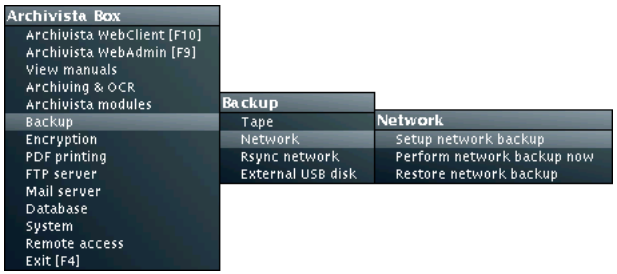
Running the command 'Restore backup' and entering the system password enables you to copy what is on the tape in the tape drive back to the hard disk of the ArchivistaBox. The more data there are the longer the process takes.


Xdialog
Volume Label:
level 0 /home/data Tue Sep 20 11:54:24 2005 afio+gzip from localho
1186+0 records in
1186+0 records out
afio: 11m+594k+1002 bytes read in 4 seconds. The operation was su
At block 1189.
OK

After the restoring process is finished a message pops up telling you that the database server was started again.

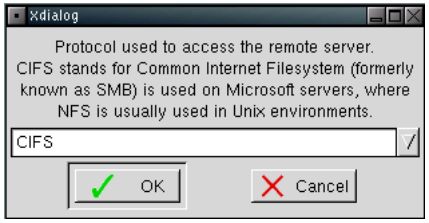
8.8.2 Network

The commands in this menu work in analogy to those for the tape backup but they let you copy the data on your ArchivistaBox to another network computer instead of an attached tape drive.



When configuring the network backup you must enter the protocol used and the IP address of the target computer.

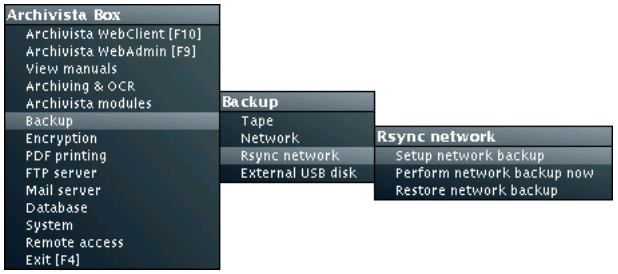
The following dialog boxes require an entry.


Xdialog
Protocol used to access the remote server.
CIFS stands for Common Internet Filesystem (formerly known as SMB) is used on Microsoft servers, where NFS is usually used in Unix environments.
CIFS
OK
Cancel

In the dialog box illustrated above you enter the IP address of the computer on which the backup copy is to be saved.

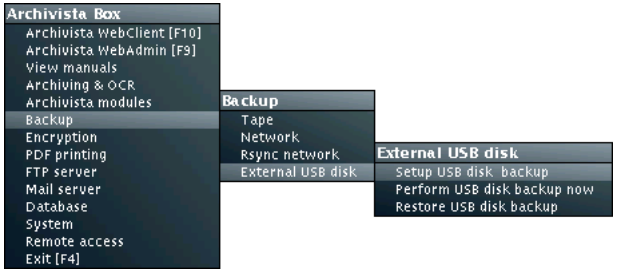
➲ There is no further confirmation. Check 'System' and 'Display system status' to see whether you have successfully configured the backup. In addition, every time the backup process has been successfully completed (usually each morning) you receive a message to that effect.

8.8.3 Rsync Network



With this menu iterm you can implement an incremental network backup. As in the backup version described above a copy of the data is written to another computer in the network but here only files that have changed are copied (and not the entire data sets).

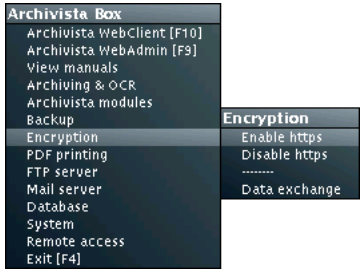
8.8.4 External USB disk



The commands in this menu work in analogy to those for the network backup but it lets you copy the data on your ArchivistaBox to an external USB harddisk instead of a network computer.

8.9 Encryption

On the one hand this menu item serves to encrypt the data streams flowing between the network and the ArchivistaBoxes PILATUS, TITLIS and EIGER, respectively. On the other hand it enables you to generate an explicit key for data exchange between two ArchivistaBoxes.



As always the individual functions are presented one after the other.

8.9.1 Enable https

Tutorial on Enable HTTPS (movie 2min) *4.18*

8.9.1.1 What is HTTPS?

Ordinarily web pages are transmitted by HTTP (= Hypertext Transfer Protocol). The data are not encrypted.

When HTTPS (= Hypertext Transfer Protocol Secure) is used the data flowing between web browser and web server are encrypted.

By running the function 'Enable https' we 'tell' the server that from now on it must encrypt all data that it sends to other computers in the network upon their request. HTTPS makes most sense for the local network and the Intranet.

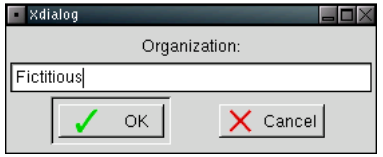
8.9.1.2 Function 'Enable https' on the server side

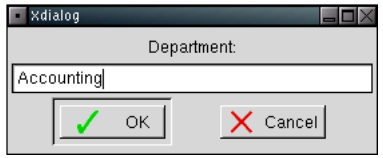
Enable https (SSL)
Please enter the system password (root user) in order to setup https (SSL) for the web server.
Password:
Cancel
Run

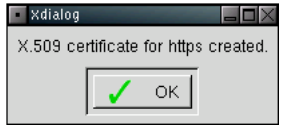
Here you must enter the IP address of the computer with which you are currently working. Remember, you are configuring HTTPS on your server just now.

The 'Country code' must be two characters long.

All further settings can be entered without any such restrictions.

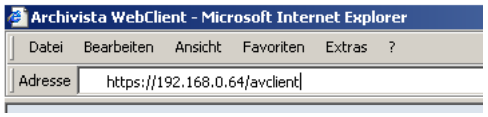






8.9.1.3 Retrieving data from a different computer in the network

After HTTPS has been configured on the server you must check whether you can access the database from any client workstation. Enter the IP address of the ArchivistaBox in the browser window but do not precede it with `http://` but with `https://`. Add `/avclient`.



Now it emerges whether your settings are correct or not. You should be able to see your data now.

➲ Please note that the ArchivistaBoxes work with the encryption certificate from Verisign. It may happen that you get warning messages with regard to specific details. If, for example, the system times of server and client workstation do not fit together, a warning message to that effect will pop up on the screen even if encryption as such is not affected. Please note the following important rule: **if you can access your data with `https://` and the IP address of your server, then your data is transmitted in encrypted form.** If you read possible warning messages closely, you will see that it is mostly unimportant partial aspects that Verisign wants you to inform of.

8.9.2 Disable https

Here you simply enter the root password. Subsequently, the HTTPS mode is turned off again.

8.9.3 Data exchange

ArchivistaBoxes utilize encryption to guarantee safe transport of data, for example when USB sticks are used as export and import media. The source PC and the target PC must be equipped with the key.

The system used here follows the principle of 'public key-encryption'. It uses two keys - a public key known to everyone and a private key known only to the recipient. The public key is used for **en**cryption, the private one for **de**cryption. When Peter wants to send a secret message to Heidi he encrypts it with Heidi's 'public key' and Heidi decrypts it with her 'private key'.

First, prepare the USB stick, then create an encryption key and put it onto the USB stick. Go through the following steps.

- Attach the USB stick to your Windows PC
- It appears as 'Removable medium' on your PC
- Right click your mouse: format as Fat32
- Right click your mouse: create a directory / folder named 'key' on your USB stick (lower case)
- Right click your mouse: 'Eject'

Then, go to your server (the 'target' ArchivistaBox) and run function 'Encryption'. (As yet without connecting the USB stick.)

Wait for the message that the key has been created: you will see the word 'done'. Should the phrase 'not enough entropy' appear, simply wait for the process to finish eventually even if it takes longer, or create entropy yourself by pressing keys at random on the keyboard.

After you see the word 'done', click 'OK' and attach the USB stick to the ArchivistaBox. The following dialog box appears:

Enter the password, click 'Run' and the key will be copied to the USB stick. (The private key remains on the 'target' ArchivistaBox.)

To convey this key to the 'source' ArchivistaBox simply insert the USB stick in this second box and the key will be copied onto it. When you now export documents from the first box they will be encrypted. And when you import them to your server or target box they can be decrypted by means of the private key.

➲ Let us assume that the ArchivistaBoxes with the data to be encrypted before the export are far away and conveying the public key to these boxes is not feasible. In this case you can send the public key to your partner by e-mail. (The 'public key' is stored in ASCII format and there are no special characters.) In such a case the file 'key' must be copied on site first to the USB stick and from there to the ArchivistaBox.

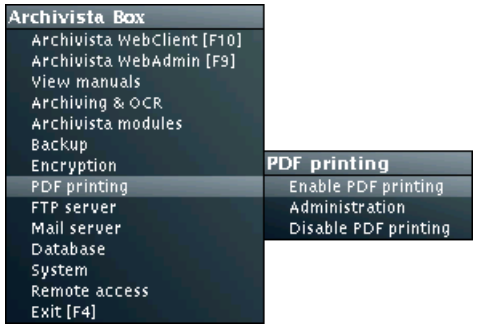
8.10 Print server

The function Print server sets up a virtual printer in order that files may be virtually printed and the resulting print spools may be archived in the ArchivistaBox.

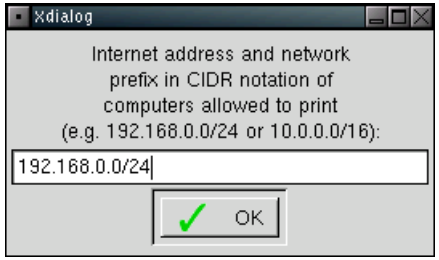
The main point is to open the ArchivistaBox for files (print spools) that users print somewhere in the network by means of the Archivista printer.

8.10.1 Enable print server

Tutorial on Enable print server (movie 2min) *4.16*



To make the ArchivistaBox susceptible to files coming from outside this function must be run. After entering the password you receive the following message:



Here you enter the network subnet from which files are accepted.

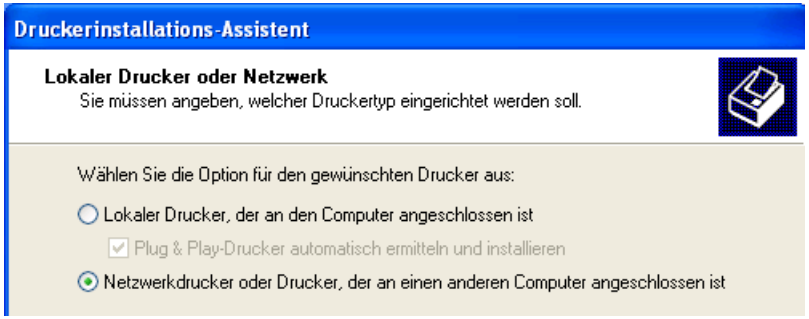
This is one side of the print server. Please note that the other side to it is configuring the Archivista printer on client workstations from which documents are sent to the ArchivistaBox.

For this purpose we need the IP address of our ArchivistaBox. By clicking 'System' and 'Display system status' we can find out this address. In our case it is 192.168.0.71.

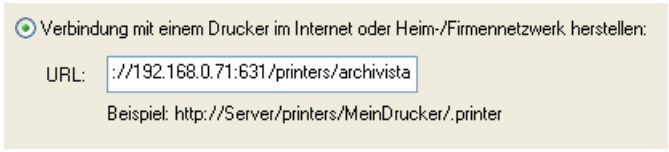
8.10.1.1 Configuring the Archivista printer

The following example stems from a German WindowsXP workstation.

Go to 'Printers' ('Drucker und Faxgeräte') and doubleclick 'Add printer' ('Drucker hinzufügen'). Click 'Proceed' ('Weiter').


Druckerinstallations-Assistent
Lokaler Drucker oder Netzwerk
Sie müssen angeben, welcher Druckertyp eingerichtet werden soll.
Wählen Sie die Option für den gewünschten Drucker aus:
Lokaler Drucker, der an den Computer angeschlossen ist
Plug & Play-Drucker automatisch ermitteln und installieren
Netzwerkdrucker oder Drucker, der an einen anderen Computer angeschlossen ist

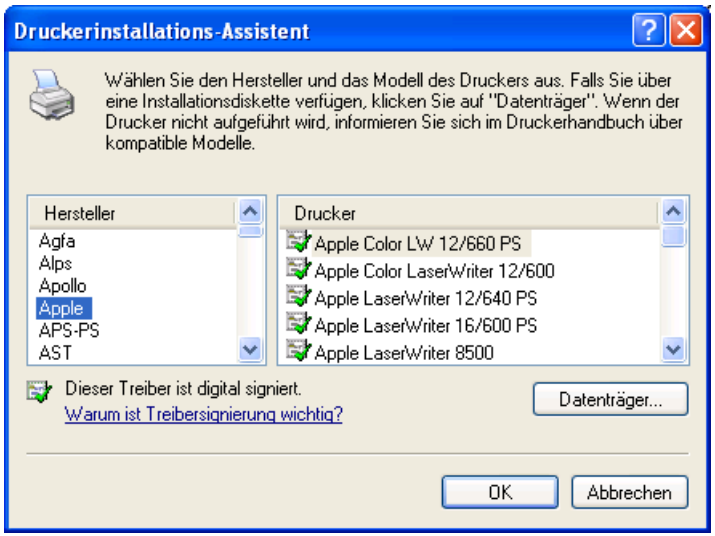
Afterwards you specify that it is a **network** printer you want to add.



In the dialog box that comes up next you enter the IP address to which the files that are to go into the archive and which have been printed with the Archivista printer should be sent.

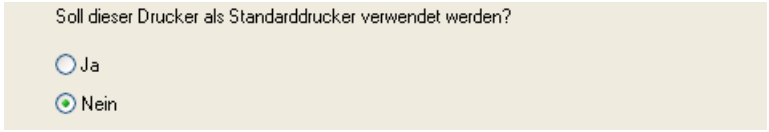
➲ In our case the entry reads: **http://**192.168.0.71**:631/printers/archivista.** In your own case the part of the entry that is not bold must show your own IP address. The first part `http://` and the last part `:631/printers/archivista` are fixed. 631 is the printer port that is used.

➲ Please note too that the last part of the entry ('archivista' in the above case) denotes the database. If you want to configure a virtual printer for a new database, 'archiv4' for example, then you make the following entry: http://192.168.0.71:631/printers/**archiv4.**



Choose the Apple LaserWriter printer. Do not get confused by the names. We are still in the middle of installing a virtual Archivista printer which, however, should take over a few properties from the Apple printer. It could also be the printer of a different manufacturer.

➲ If the printer model you are choosing has the word 'Color' in its name, then the virtual printer can print color images and you will eventually see your documents in color. If the word 'Color' is missing, then the virtual printer prints in black and white.



Say 'No' ('Nein') to the question if the printer you are installing should become the standard printer. Click 'Finish' ('Fertigstellen').

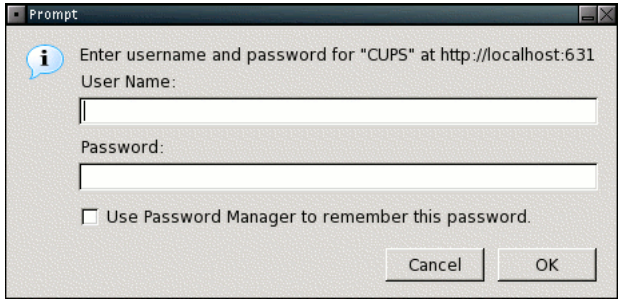
If you look again at 'Printers' you see the newly configured Archivista printer.

Those wishing to send a copy of a document from a third party application to the ArchivistaBox from now on choose 'Print' and the printer with the database name. In our case this is 'archivista'. If you need to print a document landscape fashion, then choose 'landscape' in the print dialog box.

8.10.2 Administration

When the print server is activated you can administrate this printer also directly from the ArchivistaBox.

After you have run the function the following dialog box comes up and asks you for ID and password:



Login with user 'root' and the corresponding password. You will see an overview of the existing printers.

Here you can administrate the printers and create additional printing definitions – for example to print into a second database.

You can manage printers here and open further print definitions, for example, in order to print in a second database. To do so, you have to set up an appropriate printer for each database. This means that if you want to print 'archivehello' in the database, you also require a printer with the name 'archivehello'.

➲: It is often the case that you want to print certain files in colour and others in black and white in the archive. You can then set up a printer that contains the required scan definition, separated from the database name by a hyphen. For example: You can use the `archivehello-colour` printer to print in the 'archivehello' database and the system also checks whether the 'color' scan definition exists during the import. If this is the case, the options of each scan definition become valid. In other words, the print jobs are post-processed with the extended options.

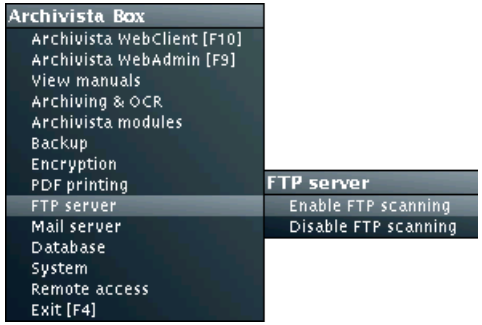
8.10.3 Disable print server

With this function you turn off the print server.

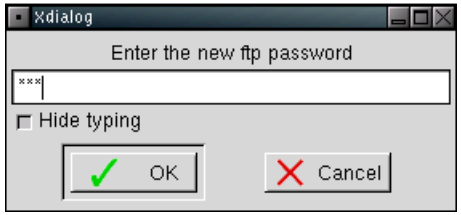
8.11 FTP Server

This function allows you to scan directly to the ArchivistaBox by means of a network scanning device.

8.11.1 Enable FTP file upload



Run the function 'Enable FTP file upload'.



As always you will first be asked for the root password. Then, you will be asked to enter a specific FTP password. You will be

prompted to enter it a second time. Now (password protected) FTP access to the ArchivistaBox is granted. You do not receive a confirmation message.

Check 'System' and 'Display system status' to review the status of the FTP server. In the illustration above you see the entry 'FTP server enabled'. Hence, everything is in place for you to scan directly into the database on the ArchivistaBox.

8.11.1.1 Upload in folder

The FTP server was originally developed to work together with FTP-capable digital copiers. However, it is now possible to send any documents to the archive via an FTP upload. The following folders exist at present:

- office: Import any files (a text extraction occurs rather than OCR recognition).
- pdf: Import scanned PDF files (OCR recognition occurs).
- tiff: Import scanned tiff multipage files (OCR recognition occurs).
- tosca: Cold import from ERP solution
- axapta: Cold import from ERP solution
- xerox: Import after Xerox copying

➲ Please note: If you send a file directly to the subfolder (e.g. Office), the file is always imported to the database **archivista.** If you want the file to be imported to a different database, you have to copy the file to the relevant subfolder (e.g. you have to send a file to `\office\archive` if you want to store it as an Office document in the 'archive? database).

8.11.1.2 Scanning by means of network scanners

There are many network scanning devices around today. The Xerox Work Centres may be mentioned as prominent examples. But also the more sophisticated copiers of other brands give this possibility.

Setting up such a network scanning device may not be a trivial task and scanning to the ArchivistaBox may not be successful at first go.

There are a couple of essential settings that must be made in the network scanner:

- for black and white scans you enter: `/tiff/archivista`
- for color scans you enter: `/pdf/archivista`

8.11.1.3 Scanning by means of AXIS Document Server

As examples may serve Xerox Work Centres or AXIS Document Servers. In the following we describe in a few short words how an AXIS Document Server is configured for scanning directly to an ArchivistaBox.

Configuring AXIS 7000 Document Server The AXIS Document Server is a device that works together with a number of scanners on the market and that turns 'ordinary' scanners into network scanners. AXIS Document Servers are very easy to use and often you can elegantly circumvent driver problems.

For a detailed description and user manual we refer to the manufacturer's homepage www.axis.com.

Summarising, we can say that the AXIS device is put into operation by following the steps below:

- At the beginning, both devices (scanner and AXIS Document Server) must be turned off

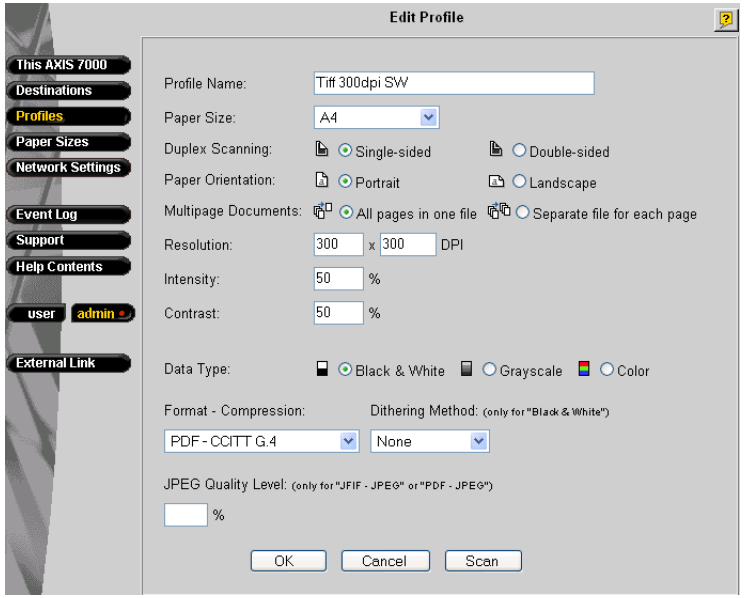
- Link AXIS Document Server with scanner (SCSI cable)
- Connect AXIS Document Server to the network
- Turn on scanner
- Turn on AXIS Document Server

Configuration and profile setting The IP address can be set manually upon start-up of the AXIS Document Server. Alternatively, the IP address can be set automatically by means of DHCP.

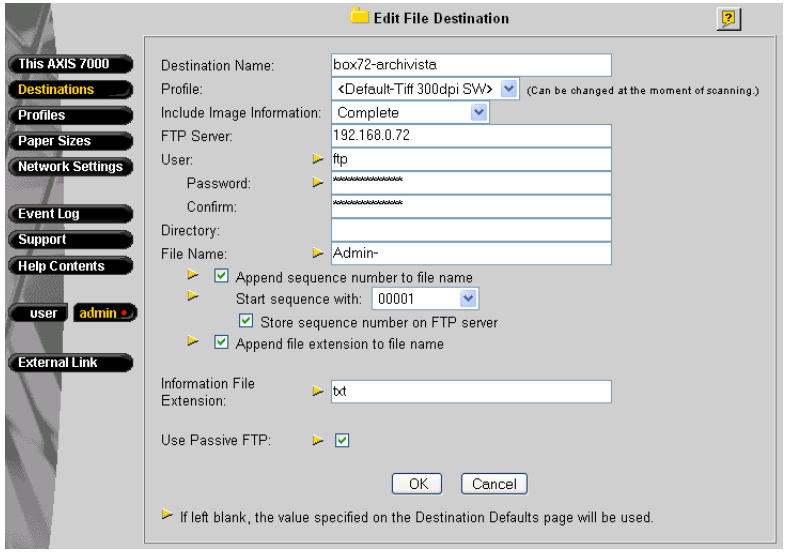
- As soon as the device is connected to the main the display shows manufacturer and version. Press 'Menu'.
- Now you can enter the IP address.
- If desired you can change to router adress and subnet mask adress by pressing the arrow keys.
- However, it is of advantage to continue configuration in a web browser on a network PC as soon as the IP address has been determined
- Confirm the setting by pressing the button for 'Select'.

For configuration through the web-based AXIS software simply enter the device's IP address in the browser window.

Under **'Profiles'** you can enter scan definitions (paper size, storage format, etc.). See below.



Under **'Destinations'** enter the IP address of the ArchivistaBox to which you want to scan. See below.



➲ **Important!** Unter 'Destinations' it is possible to enter not only the IP adress of the target ArchivistaBox but also the target database. In our case it is the default database 'archivista'. See the illustration below.

The ArchivistaBox is programmed in such a way that the characters behind the hyphen are recognised as database name.

➲ **Important!** If you put the same name for the scan definition under 'Profiles' as it exists in the WebAdmin tool of the ArchivistaBox, then Archivista 5.2 will run the following additional options specified in the scan definition (only if they are activated):

- Black/white optimization
- Default fields
- Removal of white pages
- Barcode recognition (optional module)

Illustration: The Profile Name in the AXIS document server should ideally correspond with the name of the scan definition.

➲ **Important!** Under 'Destinations' you can put the file name or the file name structure you want to allocate to the scanned documents. The information in front of the hyphen is recognised by the ArchivistaBox as 'owner'.

Naturally, this happens only if you have previously specified a user of the same name in WebAdmin. With user 'Admin' it is usually the case as it is the default user.

From the aforesaid it can be deduced that the user carrying out the scanning task can simultaneously determine the following points at the stroke of a button - i.e. by choosing the destination on the AXIS 7000 device:

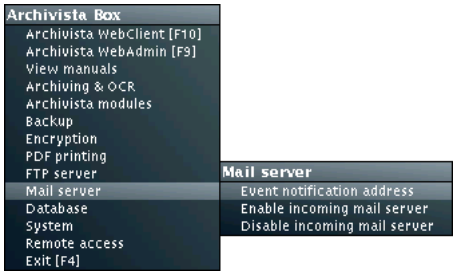
- into which database the scanner scans
- which Archivista scan definition (including options) is used
- which user is registered as 'owner'.

➲ **Important!** If your scanbox cannot make a connection with ArchivistaBox, please check if the following requirements are met: Since you are scanning via FTP the field `User` in rubric `Destinations` must show the value **ftp**. The password that you enter in the two fields that follow must fully correspond to the password with which you have enabled FTP scanning on the target ArchivistaBox. Do not be misled by the number of asterisks (*) as it will not be the same number when you log in the next time. This is a security measure in order that the password may not be guessed by the number of characters entered. Should FTP scanning still not be working, make sure that the target ArchivistaBox has FTP scanning enabled. You can check this by taking a look at the system mode (Win+F7) on the target box.

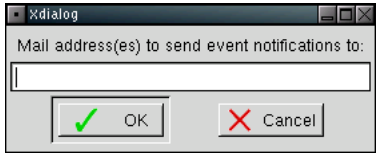
8.11.2 Disable FTP file upload

With this menu item you shut down FTP access.

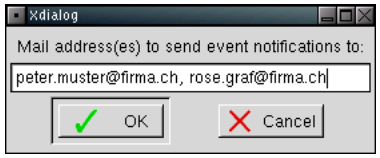
8.12 Mail server



8.12.1 Event notification address



The ArchivistaBox can send out confirmation e-mails to specified e-mail addresses. Enter here the mail addresses to which such confirmation e-mails (e.g. that backup has been successful) should be sent.



8.12.2 Activating mail archiving

You use this function to activate mail archiving for all databases. Please note that you then have to set up mail archiving in WebAdmin for each database. For more information about mail archiving, see 24.1.

8.12.3 Start mail archiving now

You can use this menu item to start mail archiving at any time.

8.12.4 Deactivating mail archiving

You can use this menu item to deactivate mail archiving again.

8.13 Database

This menu item applies only in the case of 'Titlis' and 'Eiger', i.e. when two ArchivistaBoxes form a redundant system.

8.13.1 Setup master for slave

By clicking this function you make the currently active ArchivistaBox the leading one: it contains the live configuration and the live databases and any alterations that are made affect this ArchivistaBox first.

Further down there is an example of a complete master slave configuation.

8.13.2 Enable slave mode

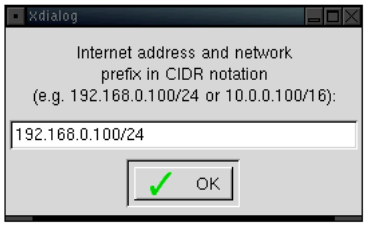
By clicking this function you make the currently active ArchivistaBox the subordinate one. The alterations made to the master box and its databases are mirrored to the subordinate box every few seconds.

The next section gives an example of a complete Master/Slave configuration.

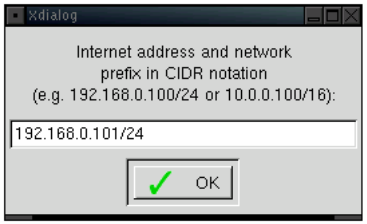
8.13.3 Example: Master/Slave configuration

8.13.3.1 Fixed IP addresses

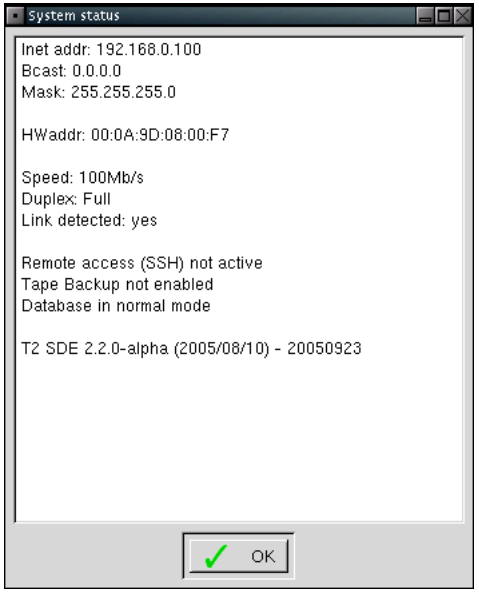
Before you can start with the Master/Slave configuration both ArchivistaBoxes must possess one fixed IP address each. (Command 'Configure network port'.)



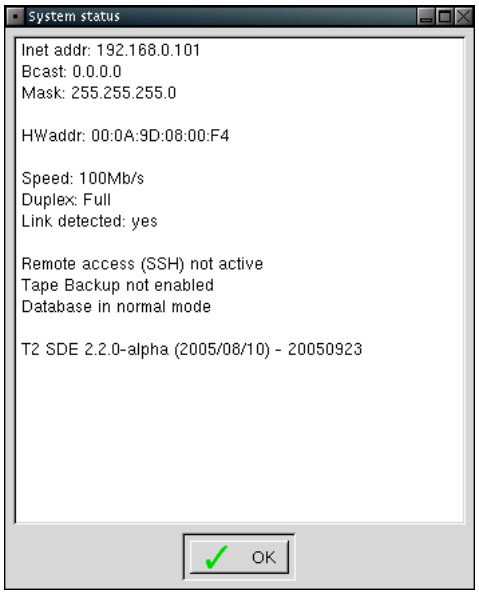
This is an example of the ArchivistaBox 'Master'.



This is an example of the ArchivistaBox 'Slave'.



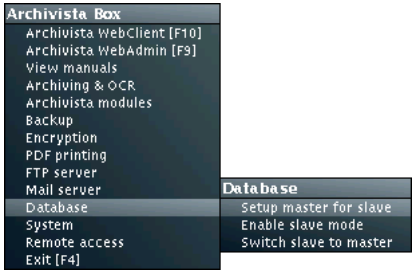
When running 'Display system status' we can see which IP address the box possesses. Above you see the 'Master'.



The above illustration shows the 'Slave'.

8.13.3.2 Configuring the 'Master'

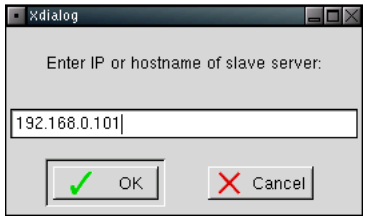
Setup the ArchivistaBox 'Master' now.



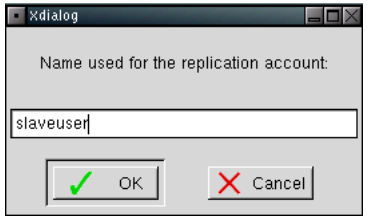
Choose the command 'Setup master for slave'.

Enable database master mode
Please enter the system password (root user) in order to enable the database master mode.
Password:
Cancel
Run

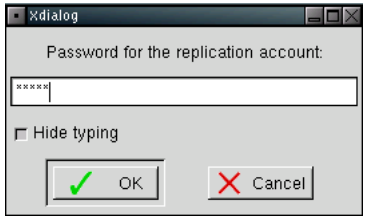
A dialog box will ask for the password.



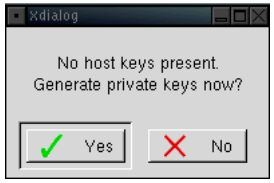
Afterwards you must enter the IP address of the ArchivistaBox **'Slave'** (without the information after the / sign). The next steps are meant to grant the 'Slave' the right to access the 'Master'.



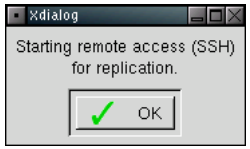
The following dialog box invites you to give the 'Slave' a name. In our example it is 'slaveuser'.



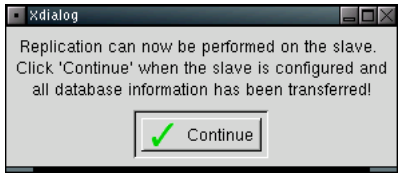
Here you allocate the 'Slave' user a password.



Unless 'remote access' was previously granted manually, the ArchivistaBox asks you now whether the keys for SSH connection should be generated. Click 'Yes'.



The following dialog box confirms the SSH connection. Click 'OK'.

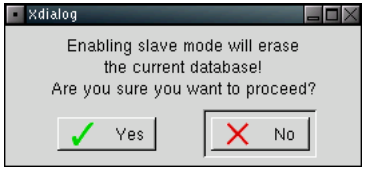

Xdialog
Replication can now be performed on the slave. Click 'Continue' when the slave is configured and all database information has been transferred!
Continue

The setup of the 'Master' box is complete. Please note: Do not click 'Continue' as yet! First you must setup the 'Slave' box.

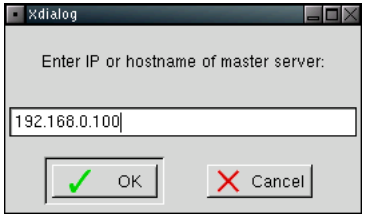
8.13.3.3 Configuring the 'Slave'

Enable database slave mode
Please enter the system password (root user) in order to enable the database slave mode.
Password:
Cancel
Run

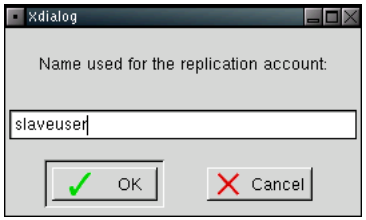
Again you will be asked for the password.



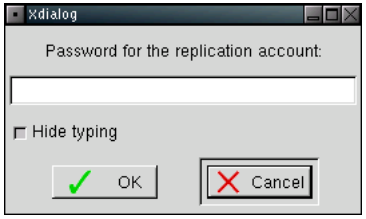
By confirming the message above you effect that existing databases are deleted from the ArchivistaBox.



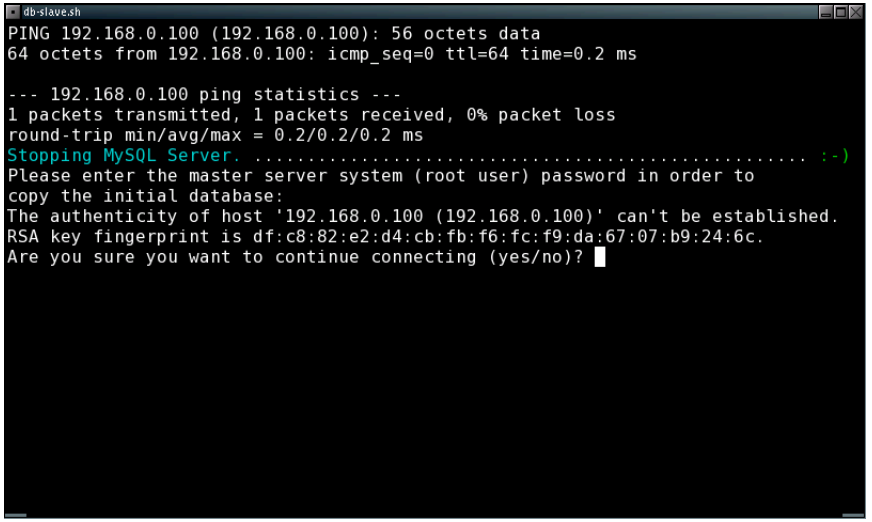
Subsequently, the fixed IP address of the **'Master'** box must be entered.



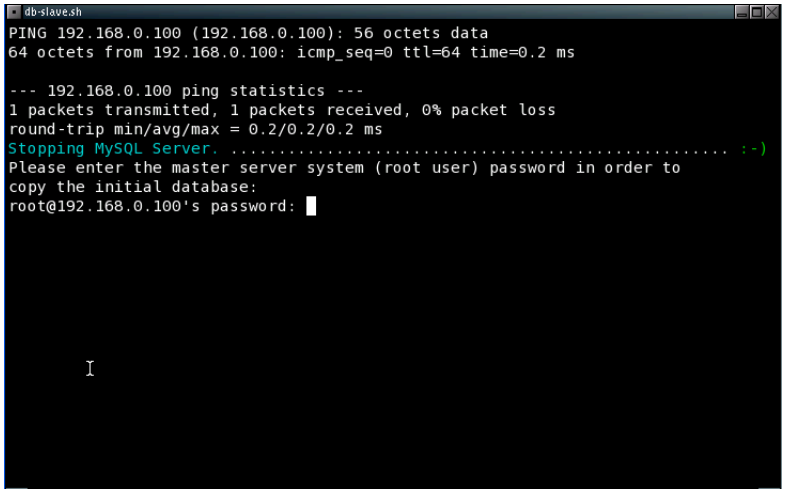
Here you must enter the name of the 'Slave' user. It goes without saying that the name must correspond to the one entered in the 'Master' box. In our case this is 'slaveuser'.



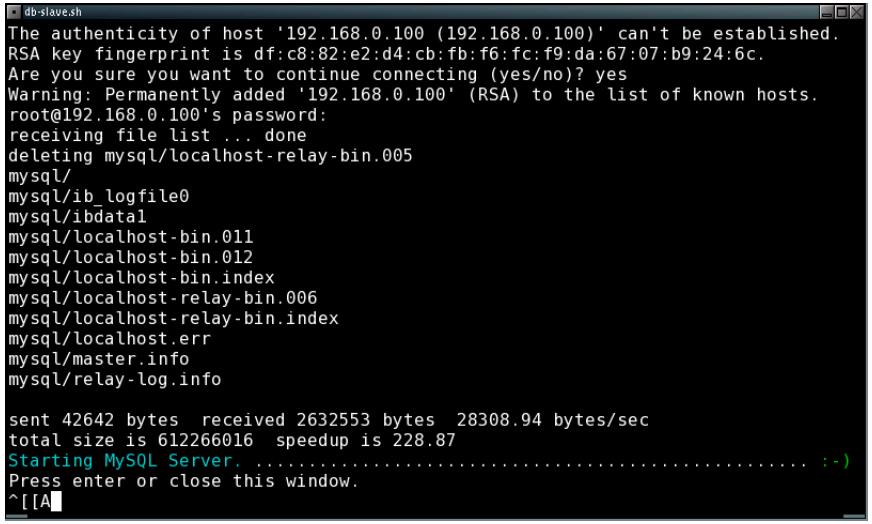
The password is the same too.


db-slave.sh
PING 192.168.0.100 (192.168.0.100): 56 octets data
64 octets from 192.168.0.100: icmp_seq=0 ttl=64 time=0.2 ms
--- 192.168.0.100 ping statistics ---
1 packets transmitted, 1 packets received, 0% packet loss
round-trip min/avg/max = 0.2/0.2/0.2 ms
Stopping MySQL Server. ..... :-)
Please enter the master server system (root user) password in order to
copy the initial database:
The authenticity of host '192.168.0.100 (192.168.0.100)' can't be established.
RSA key fingerprint is df:c8:82:e2:d4:cb:fb:f6:fc:f9:da:67:07:b9:24:6c.
Are you sure you want to continue connecting (yes/no)?

This message may or may not appear, but if it does, it should be answered by 'Yes'.


db-slave.sh
PING 192.168.0.100 (192.168.0.100): 56 octets data
64 octets from 192.168.0.100: icmp_seq=0 ttl=64 time=0.2 ms
--- 192.168.0.100 ping statistics ---
1 packets transmitted, 1 packets received, 0% packet loss
round-trip min/avg/max = 0.2/0.2/0.2 ms
Stopping MySQL Server. ..... :-)
Please enter the master server system (root user) password in order to
copy the initial database:
root@192.168.0.100's password:

Then, enter the **'Master'** password.



When you see this screen the 'Slave' is ready to take up work and to replicate all data on the box 'Master'.

Please do not forget to finalise the configuration by clicking the button 'Continue' on the box 'Master'.

8.13.3.4 Testing the configuration

Test your Master/Slave installation. Make a few alterations on the 'Master' and check whether they are mirrored on the 'Slave'.

8.13.4 Switch slave to master

This menu item serves to upgrade 'slave' to 'master' in case the 'master' fails one day.

After having run this function you can look in 'Display system status' what the current status is.

In such a case it is recommended that you order a new ArchivistaBox and connect it to the upgraded 'master' box and set it up as 'slave'. This way you have a fully redundant system once again.

During the time when you work with one ArchivistaBox only you can save the ongoing changes if need be by running an additional backup during lunchtime. See 8.8.1.2.

8.13.5 Clear master binlog

In the background of ArchivistaBox MySQL databases are at work. In order that all changes at database level can be traced so-called binary log files are created. In general we must reckon that for each piece of information added we need almost the same amount of space temporarily for the binary log files. Because ArchivistaBoxes save mostly image data, a similarly large mass of data may be created in the binary log files. For this reason you can delete these log files with the help of this menu item at regular intervals (over months/years).

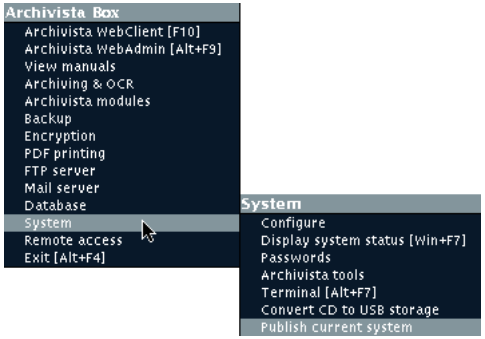
8.13.6 Clear slave binlog

This option is needed only if you work with a master/slave configuration. Underlying this option is the same principle as in the case of 'Clear master binlog'.

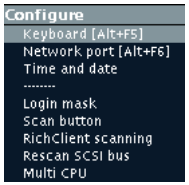
➲ After the binary log files have been deleted a check if the master/slave configuration is still intact is mandatory. In case of doubt the master/slave configuration is to be set up afresh.

8.14 System

Tutorial on Set system password (movie 2min) *4.17* **Tutorial on Unlock documents (movie 2min)** *4.17* **Tutorial on Restart OCR (movie 2min)** *4.17*



8.14.1 Configure



8.14.1.1 Keyboard (Alt+F5)

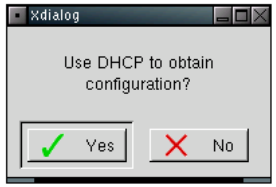
Choose the keyboard layout here. The following layouts are preset.

- US
- French
- German
- Italian
- SwissGerman
- SwissFrench
- SwissItalian

To choose a different keyboard layout click 'Other' and enter the abbreviation for your keyboard. When in doubt choose the US layout as all keyboards accept it.

8.14.1.2 Network port (F6)

Here you define the network port.



In most cases you will proceed using DHCP. Addresses will be allocated automatically. By running the next menu item you can see which addresses have actually been allotted.

Please contact your network administrator if you feel unsure about what to do with regard to the network configuration.

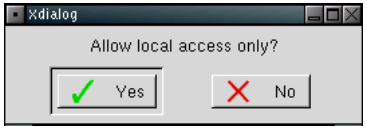
8.14.1.3 Time and date

Here you determine system date and time.

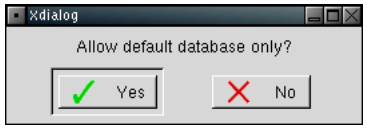
8.14.1.4 Login mask

With this menu you can tailor the login mask of the WebClient to your needs (but not that of the WebAdmin tool). Advantage: you may not have to enter the name of the host each time you log in, for example.

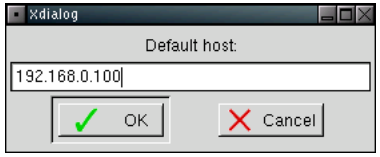
As always you will be asked for the password.



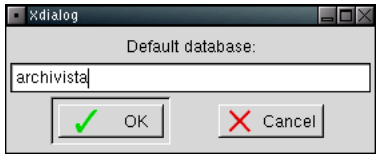
If access should be granted only from the local computer, click 'Yes'.



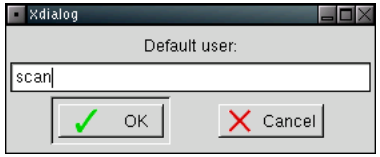
If always the same database is to be accessed, click 'Yes'.



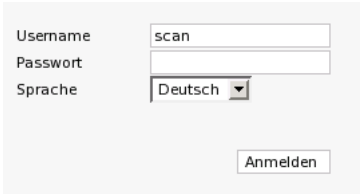
You will be asked for the host on which the database is located.



You will be asked for the name of the database to which access should be enabled.



Here enter the user whose name should appear in the login mask.



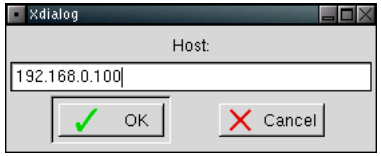
In our example the fields for host and database are no longer displayed.

8.14.1.5 Scan button

.

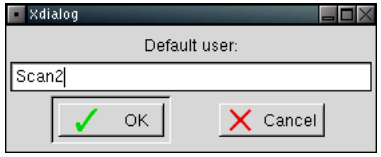
This function enables you to use a numerical keypad to start the scanning process. It makes scanning a bit easier and more comfortable.

As always you will be asked for the ArchivistaBox root password.

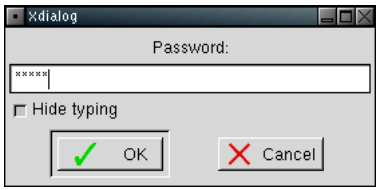


Then, enter the host on which the database lies into which you want to scan.

Here (see above) you must specify the database name.



Finally you enter the user...



...and the user password.

➲ Do consult 45.3, if you want to know more about the settings needed for remote scanning with an ArchivistaBox.

After these entries have been made there are no further obstacles to your scanning by means of a numerical keypad.

➲ If the AccessLog function is turned on and you want to use the scan button, then you can only work with it if a connection to the WebClient has been established.

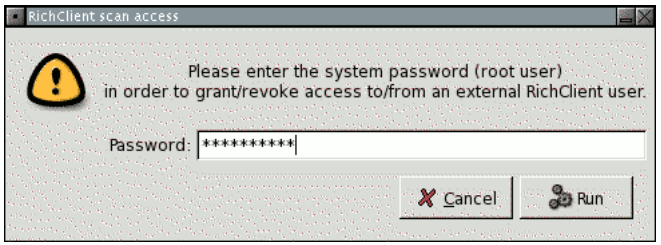
➲: You can find an LCD display under 44.10 which makes it even easier to trigger scanning.

8.14.1.6 Rescan SCSI bus

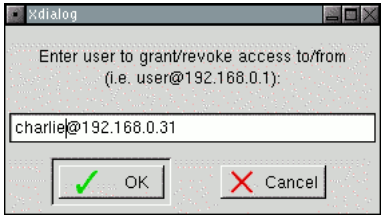
This function only plays a role if you scan with a SCSI scanner that is directly attached to an ArchivistaBox. If the users cannot scan, the reason may be that the SCSI-interface is not active (possibly because the scanner was attached to the already running ArchivistaBox system). In this case you can activate the SCSI interface by running 'Rescan SCSI bus'.

8.14.1.7 Admin user (RichClient)

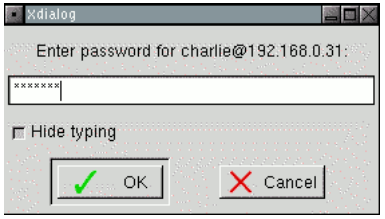
In order to be able to use an ArchivistaBox for scanning with the Archivista RichClient as described in 45.4.3 (and for some other functions) one must set up a user that has access rights. In the following a short step-by-step guide regarding setting up this user is given.


RichClient scan access
Please enter the system password (root user)
in order to grant/revoke access to/from an external RichClient user.
Password: **********
Cancel
Run

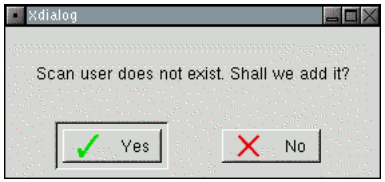
The illustration above shows that the system as always requires a password first.



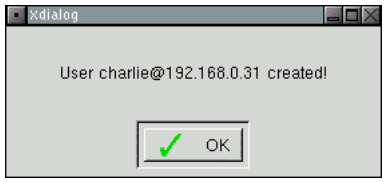
Then, a user name in the required format must be entered.



Next, a password must be assigned.



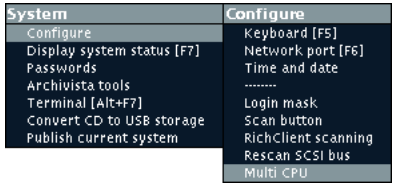
When the message above is confirmed, then a user is created – under the condition that no user of that name exists as yet. See the message below:



Now everything is in place for RichClient scanning with an ArchivistaBox used as intermediate scanning station.

8.14.1.8 Multi CPU

With this function it is possible to address mainboards with multiple core CPUs.



8.14.1.9 Configuring WebConifg

You can use this menu item to activate or deactivate the WebConfig application. Although WebConfig is activated as standard, you can deactivate it here or activate it again at a later stage.

8.14.1.10 Extended form for login

When you call the WebClient, WebAdmin or WebConfig, you will see that there are direct links to the applications and handbooks on the upper left-hand side. If you do not want these to be displayed, you can deactivate the 'Extended form for login'.

8.14.1.11 Determining the table structure for images

The ArchivistaBox usually saves the image data (which makes up approximately 90 to 95% of data) in a table. In the case of very large archives (between 50 and 100 Gbytes), it is recommended that you use this menu item to divide the tables that contain the image data into smaller units.

Use this menu item to login. The system then asks you how many folders you want to save in a table. Under normal circumstances, you can enter the value '10' here. In other words, a new table is opened automatically for all ten folders.

Of course, you can enter any value between 1 and 100. If the level of colour is high, you should keep the number of files as small as possible (e.g. 5). If you only work with black and white images, the value can be higher (e.g. 20).

➲ You can use '0' to merge tables that you have already divided into one table. This option is available for reasons of compatibility. Remember to save your data before you divide the tables. If a power cut occurs during the conversion, the tables may be lost (although this is unlikely).

8.14.2 Display system status (Win+F7)

Here you can gain information about the system. You see, for example, which network port was allocated. This is important when the ArchivistaBox is used in a network.

8.14.3 Passwords

The ArchivistaBox has administrator passwords on three levels:

- Admin
- User 'archivista'
- User 'root'

The user named 'Admin' is designed for administration tasks in the WebAdmin tool, i.e. the creation of fields and input masks or

the setting up of new users. User 'Admin' is an ordinary user with superuser (SYSOP-) rights **on the level of the database.**

User 'archivista' is designed for the 'external relations' of the ArchivistaBox. 'archivista' allows the exportation or importation of data, for instance. Internally, in the database, it corresponds to user 'SYSOP'. 'SYSOP' cannot be edited in WebAdmin.

User 'root' possesses the most far-reaching rights both within MySQL as also for the entire system. Only user 'root' may create or delete databases or make system configuration changes.

8.14.3.1 Set system (root) password

Tutorial on Set system password (movie 2min) *4.17*

Run this command and set the new system (='root') password. User 'root' has rights both on the level of the operating system as on the level of MySQL. The password is the same and is defined here.

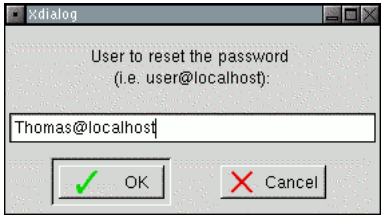
8.14.3.2 Set Archivista passwords

Tutorial on Set system password (movie 2min) *4.17*

Run this command and set the new 'archivista' password. User 'archivista' possesses the same password as user 'SYSOP'.

8.14.3.3 Reset user password

Should a user have forgotten his/her password it can be reset by means of this command.



8.14.4 Archivista tools

8.14.4.1 Unlock documents

Tutorial on Unlock documents (movie 2min) *4.17*

With this command you make locked documents available again for editing. See also 26.7.

8.14.4.2 Clear log table

By means of this submenu item you can delete the OCR log file. (For information on how to look at the log file in the first place compare section 8.5.6.2.) Removing the log items can be helpful if the text recognition module does not seem to work although it is turned on. In certain cases a 'warped' log entry may obstruct the correct functioning of the OCR software.

8.14.4.3 Remove documents or folders

With the help of this menu item you can delete certain documents or folders for good even if the documents have already been archived. Please use this command with the necessary prudence. In principle you need this menu item to delete documents definitively only after the retention period according to law has elapsed.

8.14.4.4 Restart OCR

Tutorial on Restart OCR (movie 2min) *4.17*

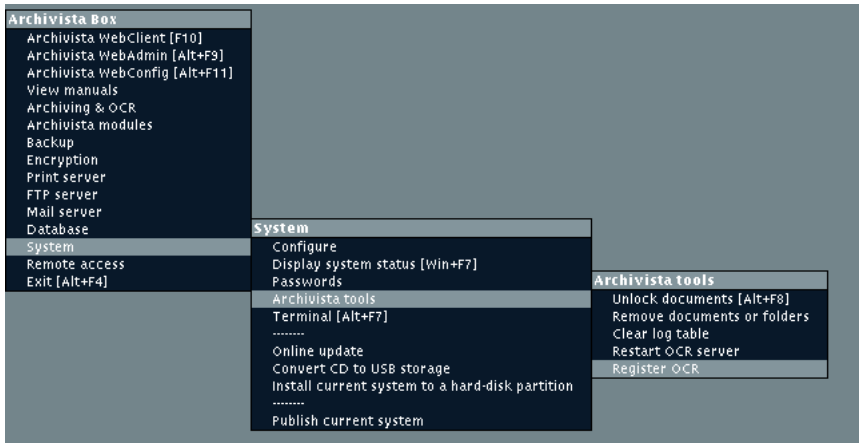
If for some reason OCR is not running, use this function to restart it.

8.14.4.5 Register OCR

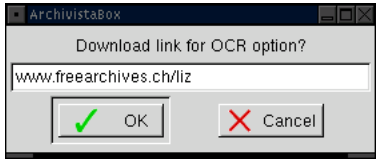
The ArchivistaBox contains various text recognition programs to ensure that the scanned pages are formatted by the text recognition software for the free text search.

A commercial OCR option is available in addition to two Open-Source text recognition programs (Tesseract and Cuneiform). Since 1.1.2007, it has been possible to obtain the OCR licenses from the freearchives.ch association (see www.freearchives.ch). You receive three sets of information: a download link, a file name and a password. You then have to register the OCR option to activate it.

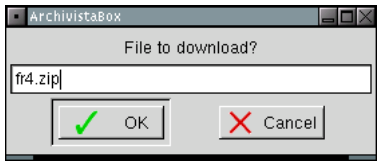
Under 'System' in the system menu, select the menu item 'Archivista tools' and then select 'Register OCR':



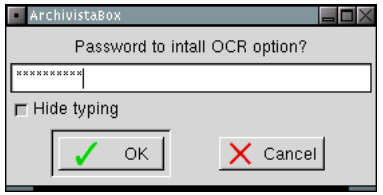
The following window is then displayed



Enter the download link that you received (this should be **www.freearchives.ch/liz**). Click on *OK* and enter the file name that you received for the download (e.g. fr41.zip):

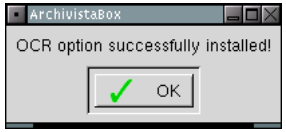


Confirm your entry with *OK* and enter the password that you received:



Once again, you confirm your entry by clicking *OK*. The download is now activated and the following is displayed on your screen

Once the download is complete, the following confirmation is displayed on your screen and you close it by clicking *OK*.



The OCR option is then active and can be used immediately.

8.14.5 Terminal (Alt+F7)

This menu item is reserved exclusively to the developers of the ArchivistaBox.

8.14.6 Online update

You can use this menu item to update the ArchivistaBox to the latest version directly via the Internet. After you call this menu

item, you have to enter the root password. You then have to enter the required version number.

This number has the following structure `YYYYMMDD`. For example, to install the release from 17.5.2009, you would enter`20090517`.

The current ISO file (approx. 700 Mbytes)is then downloaded. This process can take some time depending on your connection. Once the download has been completed successfully, the update is carried out on the alternative system partition. The system then issues a confirmation message to inform you that the update has been carried out successfully.

➲ Please note that the system connects to the server automatically. Unfortunately, the reaction time of download servers can, occasionally, be extremely slow. If this is the case, you can restart the ArchivistaBox while the update is in progress (so long as the download is successful) and start the update again. If the download is unsuccessful again, carry out an update in accordance with 8.14.9.

8.14.7 Publish current system (installed ArchivistaBox)

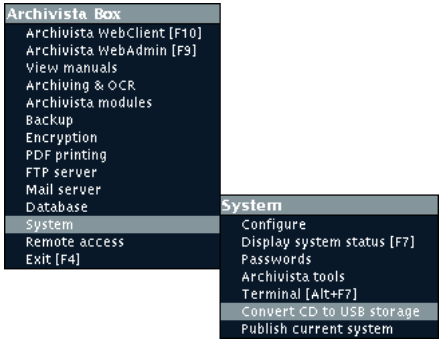
This menu item is used to copy an existing ArchivistaBox (incl. one or more databases) to a DVD or to an external hard disk, respectively. How archives are published is described in 38.

8.14.8 Convert CD to USB storage

With the help of this function an installation USB stick may be copied from an installation CD. (See also 8.14.9)

This is advantageous in all cases where you want to install ArchivistaBoxes that do **not** possess a CD drive. For the creation of the USB stick you do need a box that contains a CD drive, though.

Insert the installation CD and run the command.

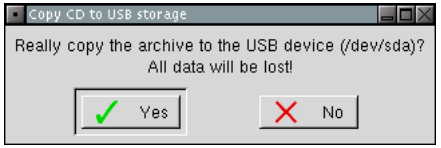


Enter the root password.

Copy CD to USB storage
Please enter the system password (root user) in order to copy a CD to USB storage device.
Password:
Cancel
Run

Attach the USB stick to the device...

...and confirm that the writing process can be started.



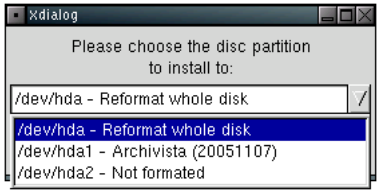
After the copying process is finished a message to that effect is displayed. The USB stick may be removed now.

8.14.9 Install current system to a hard-disk partition (only live CD

➲ Compare also 44.4 for the concrete steps when installing the system afresh.

If this menu item appears at all, then you are working with the 'live-CD' or the 'live-USB stick'. (If it does not appear, you are working from the hard disk.) The 'live CD' is a CD which, when run, gives you the same look and feel as the live system but which can also be used for installation. The 'live-USB stick' is the same as the 'live-CD' but on a different medium.

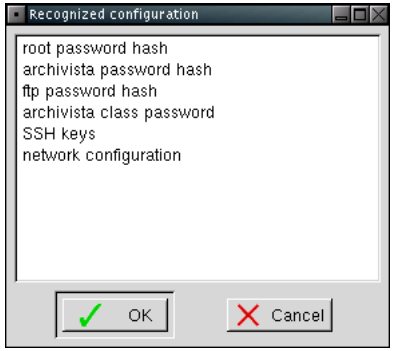
Running this command means that the system is installed on the hard disk of the ArchivistaBox. You can choose whether you want to include the entire hard disk and thereby delete existing data or if the new system is installed while existing data are kept.



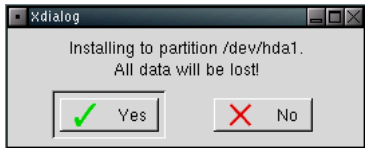
To format the entire hard disk and install the program run the first submenu item. The second submenu item shows that there is an Archivista installation dated 7 Nov 2005 on one partition. To overwrite that program version but keep the data you opt for the second submenu item. If you are still unsure whether you want to update your archive to the new program version you can choose the third possibility and install the program to the as yet unformatted partition.

The following refers to the second alternative: update of the existing archive to the new program version. Here you can even choose if not only the data should be kept (existing documents in the database and database settings, etc) but also the configuration (passwords, IP address, etc).

By clicking `/dev/hda1␣-␣Archivista␣(20051107)` also the configuration is kept.



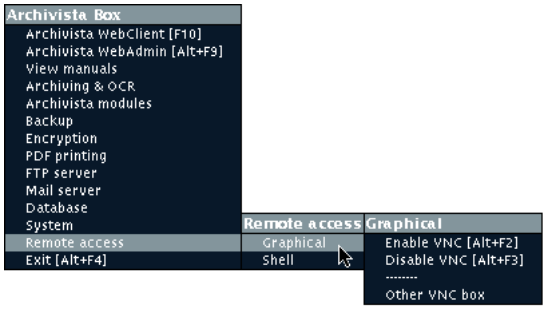
You receive a short overview of which configuration data are transferred to the new installation.



➲ Please note that also with this option you will be warned that 'all data will be lost' but this refers only to the old program version.

8.15 Remote Access

8.15.1 Graphical

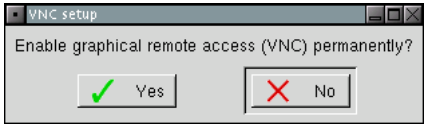


8.15.1.1 Enable VNC (Alt+F2)

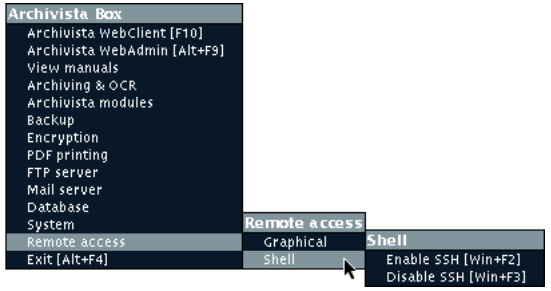
Tutorial on Enable VNC (movie 2min) *4.15*

With 'Enable VNC' you determine that the ArchivistaBox with which you are currently working can be accessed from other computers in the network by means of VNC. 'Disable VNC (F3)' will close the graphical remote access again. The command 'Other VNC box' serves to 'remotely enable remote access' so to speak.

Remote access is a maintenance account and should not be open normally. However, it is possible to keep it open permanently in special cases. Simply choose the corresponding button.



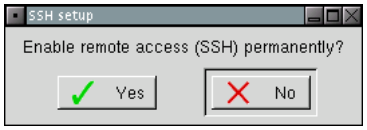
8.15.2 Shell



Tutorial on Enable SSH (movie 2min) *4.14*

For a secure connection to the console you run 'Enable SSH (Alt+F2)'. With 'Disable SSH (Alt+F3)' you close the connection again.

Remote access is a maintenance account and should not be open normally. However, it is possible to keep it open permanently in special cases. Simply choose the corresponding button.



8.16 Exit (Alt+F4)

With this command you shutdown the ArchivistaBox.

8.17 Function keys

The most important functions of the ArchivistaBox can be accessed by function keys:

F1	Help
Alt+F1	Documentation in English
Win+F1	Documentation in German
Alt+F2	Enable remote access
Alt+F3	Disable remote access
Alt+F4	Shutdown ArchivistaBox
Alt+F5	Configure keyboard
F6	Configure network port
Win+F7	Display system status
Alt+F7	Open terminal window
F8	Unlock documents
Alt+F9	Start WebAdmin
F10	Start WebClient
F11	Turn on/off full screen mode
F12	Quit WebClient/WebAdmin

Part IV

WebClient

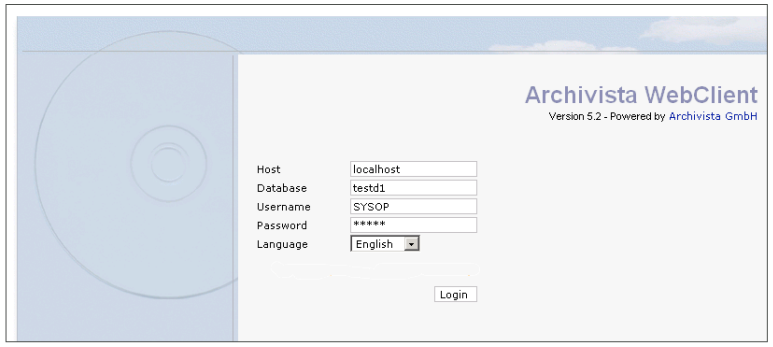
9 User Manual WebClient 2009/I

9.1 Login

Tutorial on WebClient Login (movie 2min) *4.5*

Access to the Archivista database takes place via the Internet browser. Enter Host, Datenbank, User-ID and Password. This information will be given to you by your system administrator.

➲ Note: It is possible that you only have to enter a user name and password. It is also possible that you may be able to enter the database name but not the name of the computer. Initially, all four values are required, although the first three have default values (localhost, archivista, Admin). The default password is 'archivista'.

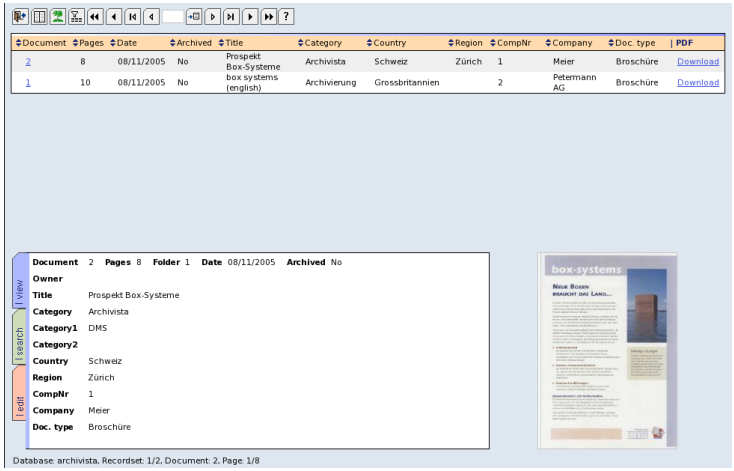


9.2 Views

There is a main view as well as a page view.

9.2.1 Main view

The main view consists of a list of documents (upper half of screen), details of the currently active document (lower half of screen on the left) and the image of the currently active page (bottom right).

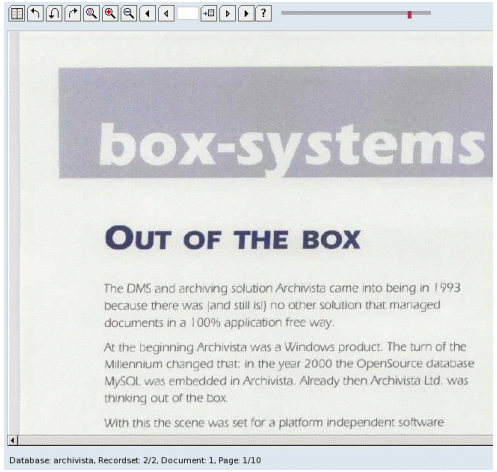


The status bar at the bottom left shows information on the currently active document: which document of how many it is, what number the document possesses and which page is displayed.

Depending on which version of the WebClient you possess, you cannot only search documents and pages but you can also edit and delete them. To the left of the document detail area you find tabs in different colors. To find out more information about searching please go to 11.2. To find out more about the editing mode see 11.3.

9.2.2 Page view

Page view gives you the best possible view of the currently active page. The image is automatically scaled to your browser window.


box-systems
OUT OF THE BOX
The DMS and archiving solution Archivista came into being in 1993 because there was (and still is!) no other solution that managed documents in a 100% application free way.
At the beginning Archivista was a Windows product. The turn of the Millennium changed that: in the year 2000 the OpenSource database MySQL was embedded in Archivista. Already then Archivista Ltd. was thinking out of the box.
With this the scene was set for a platform independent software
Database: archivista, Recordset: 2/2, Document: 1, Page: 1/10

10 Navigation

Navigation, i.e. moving from one document to the next and 'leafing' through pages of individual documents, is not done by commands from the menu but by clicking the relevant icons.

10.1 Main view

Page view: display pages (Alt+v). With this command you change to a view which shows the active page scaled to your browser.

➲Please note that you can also change between main and page view by clicking the image once.

Quit program (Alt+q)

Select all records (Alt+a)

Browse backwards (Alt+f). This command refers to the documents of the current selection. With this icon you can scroll backwards through the document list.

View previous document (Alt+p). This command refers to the documents of the current selection.

View first page (Alt+1). This command refers to the pages of the currently active document.

View previous page (Alt+2). This command refers to the pages of the currently active document.

Go to page (Alt+g). Before clicking the icon please enter a page number in the field to the left.

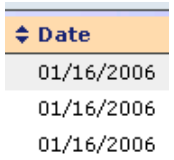
View next page (Alt+3). This command refers to the pages of the currently active document.

View last page (Alt+4). This command refers to the pages of the currently active document.

Next record (Alt+n). This command refers to the documents of the current selection.

Browse forwards (Alt+l). This command refers to the documents of the current selection. With this icon you can scroll forwards through the list of documents.

10.2 Sorting



The document list can be sorted by clicking on the triangles in front of the relevant term in the red bar.

10.3 Page view

Back to main view (Alt+v). With this command you change back to the main view.

Please note that you can also change between main and page view by clicking the image once.

Rotate graphic to the left (Alt+l).

Rotate graphic to the right (Alt+r).

Put graphic on its head (Alt+u).

View entire page (Alt+s).

Zoom out (Alt+w)

Zoom in (Alt+x).

Previous record (Alt+p). This command refers to the documents of the current selection.

View previous page (Alt+2). This command refers to the pages of the currently active document.

Go to page (Alt+g). Before clicking the icon please enter a page number in the field to the left.

View next page (Alt+3). This command refers to the pages of the currently active document

Next document in the list (Alt+n). This command refers to the documents of the current selection.

Continuously adjustable zooming: by moving the little red bar the page display can be easily adjusted.

11 View, Search and Edit

In addition to the main view and the page view, there are three further operating modes in the main view. Documents can be viewed, searched and edited. When you view the current document, it is displayed on the lower right-hand side.

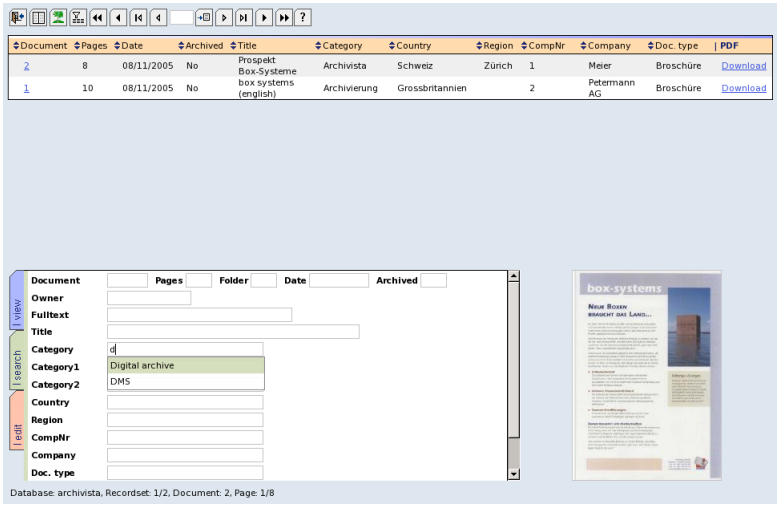
11.1 View (Ctrl+F5)

You use the blue 'View' tab in the detail view of the current file to access the view (in the middle of the area on the left-hand side).

11.2 Search

Tutorial on Searching (movie 2min) *4.8* **Tutorial on Searching (movie 2min)** *4.9*

Searching by means of the WebClient is very comfortable: there is a special search dialog box at your disposal. Simply click the green tab of the detail view of the currently active document. (You find it in the lower half and on the very left of your screen.)



You see a dialog box with empty fields - ready for you to enter your search criteria. For those fields for which pull down menus exist, you can select your search criteria by picking them from the list.

11.2.1 Full text search

If you do not want to search for your documents and pages by making use of keywords in the fields but want to search the entire text of all the pages, then use the full text search possibility.

The little window serving this special purpose is to be found in the upper right corner of the search dialog box.
Please note the following features of the full text search:

- There is no distinction between upper and lower case

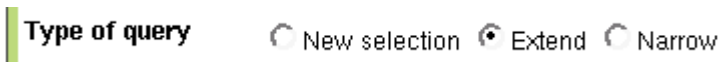
- You can look for parts of words by adding an asterisk to a word; `man*`, for example, finds 'mandarin' as well as 'manager'
- An asterisk at the beginning of the search term is ignored; `*man` does not find 'Hopeman'
- A Boolean 'AND' can be achieved by first entering the one and then the other search term, e.g. `word␣writing`. This search returns all pages on which 'word' and 'writing' are to be found
- You can enter expressions in between quotes; `"Berlin␣Wall"` finds all pages on which exactly this expression appears.

➲ The full text search can only come up with results if the documents and pages are actually available in full text. If in doubt, please ask your system administrator.

11.2.2 Extending results or narrowing them down, new selection

11.2.2.1 Option 'extend'

Multilevel searches are possible by means of the options towards the bottom of the dialog box.



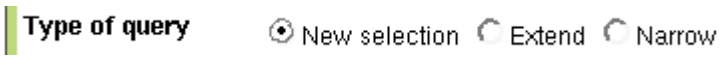
When you do a second search and you tick the option **Extend** you will see the results of both the first and the second search. For instance: I click **New selection** and choose `'Sample␣Inc.` in the field 'Client'. Then I click the **Search** button. After the results have been displayed I go back to the dialog box, I click `Sample␣Holding` and in addition I click **Extend.** Then I click the **Search** button. This two-level search yields all documents that belong to Sample Inc. AND Sample Holding.

11.2.2.2 Option 'narrow'

Narrowing down my results works similarly. For instance: I click **New selection**, choose `Sample␣Inc.` in the field 'Client' and press the **Search** button. After the results have come up I decide that I want to refine my search. I go back to the search dialog box and click **Narrow.** In the field 'document type' I choose `correspondence` and click the **Search** button. Now the system returns only the correspondence for Sample Inc. (All other document types like invoices, delivery notes, etc. are ignored.)

11.2.2.3 New selection

As mentioned above tick **New selection** whenever you want to search the entire document base for which you possess rights.



As a rule, when you start up the program **all** documents appear that you are allowed to see. 'New selection' is the default value in the search dialog box. For workflow implementations it may make sense to display a selection that has already been narrowed down (only those documents are shown, for example, which need to be processed). In such a case please refer to the specific workflow documentation.

11.2.3 Date range

The format when searching for a date range is: d.m.yy-d.m.yy, for example 1.9.03-30.9.03. This stands for September 2003.

11.2.4 Entering ranges for number fields

You can also carry out a range search for number fields. This includes the File, Pages and Folder fields. For example: By entering 10-11 in the 'Pages' field, you select all files that have 10 to 11 pages.

11.2.5 Searching with larger/smaller

When you carry out searches in date and number fields, you can also use < for smaller or equal to and > for larger or equal to.

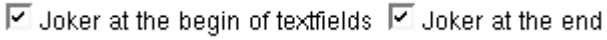
11.2.6 Searching text fields for different files

By including '!' at the start of your entry, you can search text fields for different files. For example, if you enter '!Meier' for a company name, the system finds all files that do not relate to 'Meier'.

➲: If you work with wildcards, these are also used when searching for different files. For example, if you enter '!tree', all hits that contain 'pear tree', 'tree trunk' or 'tree' are excluded.

11.2.7 Joker

At the very bottom of the search dialog box you can specify whether search strings entered in text fields should begin or end with a wildcard or joker symbol.



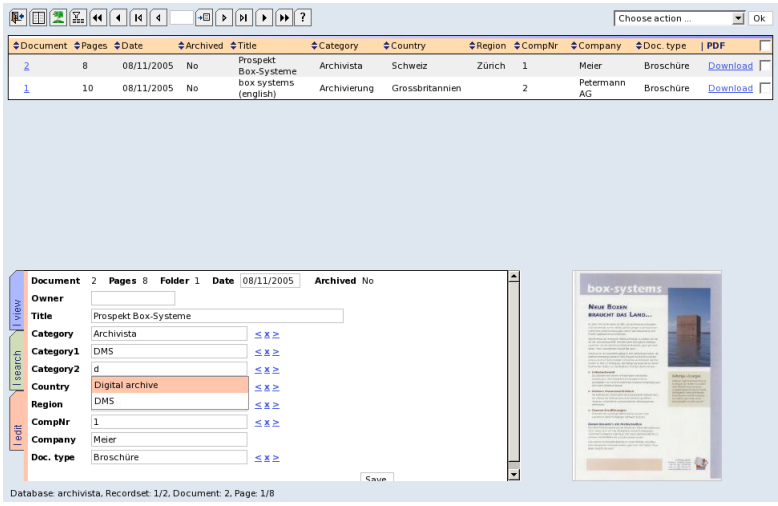
The use of this function is best explained by examples. Hence, when the option **Joker at the beginning of textfields** is ticked, then the search term `tree` returns also documents with 'Portree' or 'Crabtree' in the field in question.

When the option **Joker at the end of text fields** is activated, then the search term `tree` returns also document with words such as 'Treetop', 'Treehouse' or 'Trees'.

➲ Please note that this option does not apply to the full text search field. For the variations possible in the full text search please refer to 11.2.1.

11.3 Edit mode

There is a special dialog box in the WebClient for the 'edit mode'. Click the red tab in the detail view of the currently active document. (You find it in the lower half and on the very left of your screen.)



You see the edit dialog box specially adapted to your and your company's needs. Depending on user access rights and work sphere some fields may be locked for you while others need to be filled in with keywords. In addition, there are different field types. Below we give a quick overview.

Do not get confused, however. There is no need to learn what follows by heart. As a rule and to ensure stringent classification of documents the system guides and supports the individual user in the process of keyword allocation. The more data in a database

the more important the streamlining of possible entries. If at all unsure as to how to allocate keywords, do contact your system administrator.
As mentioned above we give a quick overview of what kinds of entry possibilities the system typically provides. Should you want to know more we refer you to chapter 36.3 and 36.6.

11.3.1 Uncontrolled keywording

You are free to enter anything you like.

11.3.2 Field types: number field and date field

The field requires a number (no text) or numbers in date format (mm/dd/yyyy or similar).

11.3.3 Controlled vocabulary

You needs must choose from a given number of entries. As soon as you make an entry, a drop down list appears from which you can pick the value you want.

Users who are allowed to add to the pre-defined entries see the following characters next to the field

< x >

By clicking > on the right hand side a field opens up and **new expressions** can be added. A field opens up.

Enter the new value, click the button 'Add' and the new expression is saved. If you want to stop adding new expressions, click < and the field closes again.

To delete controlled vocabulary, select the expression you have in mind and then delete it by clicking X.

11.3.4 Field type 'hierarchical'

Hierarchical fields allow you to create keyword trees - similar to the 'folder within the folder' principle of the Windows Explorer. Typically hierarchical fields appear in the guise of pull down menus with controlled vocabulary.

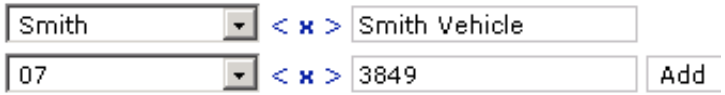
11.3.5 Field type 'Multi'

The field type 'Multi' enables you to add multiple keywords per category. You can allocate one document to several customers, for example. Multi fields typically appear together with controlled vocabulary.

11.3.6 Field type 'linked field'

This field type occurs often together with barcode processing: a text field is linked to a number field and when the latter is filled in after the barcode has been 'read' by the system, the former can be completed accordingly. But also for manual entries a linked field can make sense because unambiguousness is often only achieved with the help of numerals.
➲ Example: a recruitment company possesses several hundred customers of which some have very similar names. 'Smith Vehicle Construction Ltd.' und 'Smith Vehicle Parts Ltd. ' can be distinguished more easily when the client number appears alongside the client name.
Also this field type occurs together with controlled vocabulary. When an expression from a text field is selected the value in the number field jumps automatically to the number that is linked to the expression. And when a number is selected, the text in the text field changes to the expression that is tied to the number.

When a new expression is added clicking the X makes two fields open up:



Beside the combination 'text field'-'number code' there is also the combination 'text field'-'text code' for those cases where specific abbreviations are used in a company but where the abbreviations are not known well enough to exclude confusion. The combination 'Reporting and Consolidation'-'RC' elegantly enables correct keywording, for example. For the rest, the procedure is the same as with 'text field'-'number code'.

11.3.7 Navigation

To hop fast and easily from one field to the next in the edit dialog box you best use the tab key with your left hand.

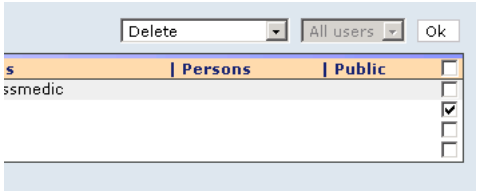
11.3.8 Document scanning

This is a special case of the Editing function insofar as an additional device (scanner) is needed and it is treated in a separate section. See 12.2.

11.3.9 Deleting

While documents are edited in the edit dialog box, the 'delete function' is run from field 'Choose action...' in the upper right corner of the screen.

First go to the Edit mode by clicking on the red tab. Thereon the pull-down field 'Choose action...' will appear in the upper red corner. Select 'Delete' and click the little box(es) beside the document(s) you want to erase.

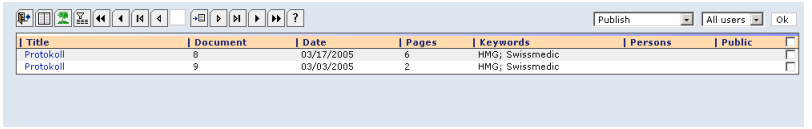


Click OK.

A dialog box will pop up to ask you whether you want to go ahead with the deleting process. Click OK again and deletion takes effect.

11.3.10 Publish and Unpublish

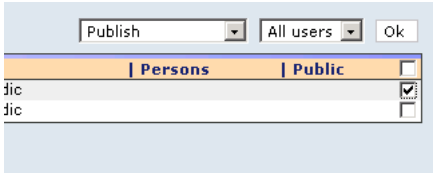
This function enables an owner of a document to make it available to an extended circle of users but without losing ownership of the document him- or herself.



In our example user 'freuler' clicks the red edit mode tab. In the upper right corner of the screen there appears a field with different actions. To publish documents to further users the action 'Publish' must be selected.

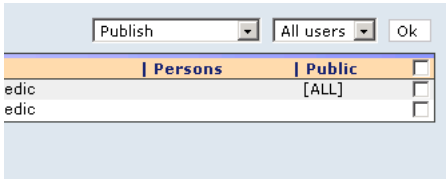
In a second step the user(s) for which the documents should be published are specified in the field adjacent to 'Actions'.

In a third step the document(s) in question are ticked.



➲ By ticking the small box at the top all documents of the current selection can be selected for the action that follows.

By clicking the button OK the action is carried out.

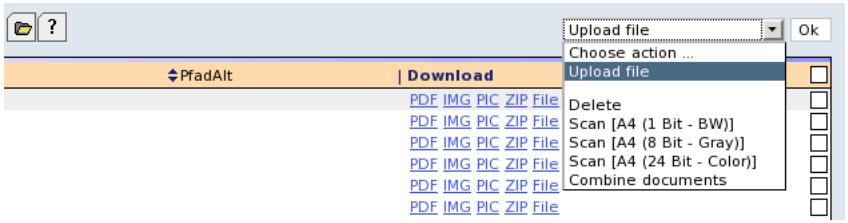


In our example one document was released for all users. This can be seen from the entry [ALL] next to the record concerned.

To reverse the publishing process choose **Unpublish** instead of **Publish.**

12 Extended functions

12.1 Uploading a file



You use this menu item (or optionally the icon shown on the left) to directly upload a file to the active database within the WebClient. After the call, the toolbar is displayed in a slightly different way.

You can use the first button to determine the file that you want to send to the archive. To the right of this, you can choose whether the file is to be processed with 'Black/White', 'Gray' or 'Colour'. On the far right, you determine whether OCR text recognition is to be carried out and if so, what type.

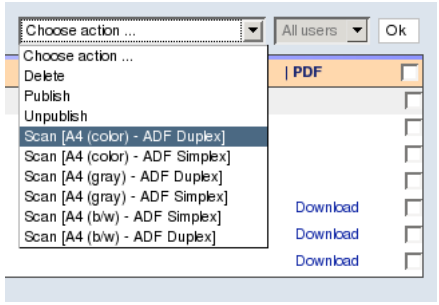
➲: At present, you can upload PDF files as well as established Bitmap files (tif, png, gif and jpg). As of version 2009/I, you can also upload all Office files. When you do this, these files are scanned automatically. The original file is also archived and you can check it out again later.

12.2 Creating and scanning documents

Tutorial on creating and scanning documents (movie 2min) *4.6*

It is the Edit mode that enables you to create documents and to scan pages.

First go to the Edit mode by clicking on the red tab. Thereon the pull-down field 'Choose action...' will appear in the upper red corner. Select 'Scan' and click the button 'OK'.

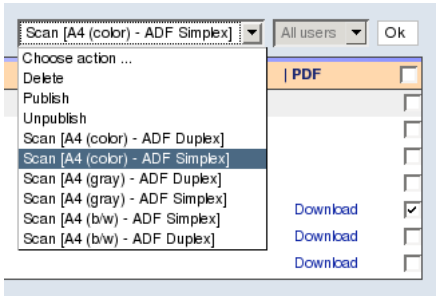


After about 30 seconds the newly created document and the scanned pages will appear. Click

or press Alt+a to refresh the screen and to see whether the pages are already there.

To add new pages to an **already existing** document follow the steps below:

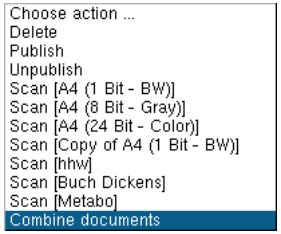
- go to the Edit mode by clicking on the red tab
- select 'Scan' in the pull-down menu 'Choose action...' in the upper right corner of the screen
- tick the document to which you want to add pages



- click 'OK'
- confirm the pop-up message.

12.2.1 Combine documents

To add a new page to an existing document one can normally go to the document in question and scan the new page. With a digital copier this does not work, however. A separate document is created for the new page that has been scanned. To move this page to the target document you need the function **Combine documents**. This function can be found in the field **Choose action** appearing in the top right corner when you go to Edit mode.



➲ If you want to combine two documents none of the two documents may have the status 'archived'.

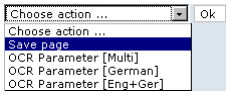
To combine two documents you must proceed in the following manner:

1. Go to Edit mode
2. Tick the two documents that you intend to combine. The document to which the other one will be copied is underlaid with gray. You can change the target document by doubleclicking the target document.
3. Go to the **Choose action** field, select `Combine documents` and click `Ok`

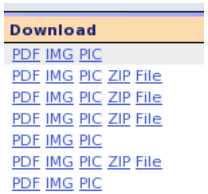
12.3 Editing and deleting pages

Tutorial on Editing and deleting pages (movie 2min) *4.7*

You can edit the pages of a document: it is possible to rotate and save individual pages, to delete them or to assign specific OCR definitions to them. All of these functions are carried out by means of the 'Choose action' menu in Page view.



12.4 Download



There are a few links on the right-hand side of the table in the main view of the WebClient. You can use these to obtain the respective document or a specific page as a PDF file or as an image file.

At present, you can download the following formats:

- PDF: Page/file as a PDF file. Depending on the option 'Convert whole file to a PDF file', the system provides the current page or the whole file as a PDF file.

- IMG: The current page is downloaded in high resolution.
- PIC: The current page is downloaded in reduced resolution.
- ZIP: An original file (source) that can be downloaded as a zipped file exists for the respective file.
- Mail: The current file was archived as a mail message. You can use this command to return the file to the mail server.
- Datei: The current file was archived as an Office file. By clicking on the link, you can restore the original file in the hard drive or open it directly.

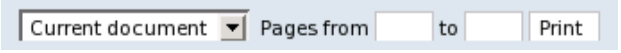
➲ If the 'Create PDF files' option is activated in WebAdmin, the PDF file can be downloaded directly. If this is not the case, the respective PDF file is created 'on the fly'.

12.5 Printing pages

In principle you can print individual pages by going to 'Page view', displaying the page in adequate size and running the print function from the web browser menu. For individual pages this is workable, however, if you need to print one or several documents in the WebClient, then it is inconvenient.

➲ This option must be activated by the administrator. label-webprinting

If the administrator has activated the option and thus enabled the printing function for you, you see at the bottom right of your screen a couple of fields, which you can use to print one or several documents.



➲ You can print more than one document only if you possess 'Editing' rights. If you do, you can select the documents in question and run the option 'Selected document (s)' under 'Current document' to print them all at once.

12.6 Barcode processing

If your Archivista WebClient 2009/I has been equipped with the barcode processing module, then you can easily scan your vouchers, invoices or reports and keywords are added automatically.

Simply insert the documents with the barcodes on them into the ADF of the scanner and scan them. Subsequently, the system will read the barcodes, decode them and write the data into the fields where they belong.

Part V

WebAdmin

13 Login and Logout

Tutorial on Login WebAdmin (movie 2min) *4.10*

Welcome to the administration of Archivista 2009/I by means of the Archivista web interface! In the following we guide you through the different menus and options. To adjust Archivista 2009/I to your needs you can set up users, create new fields and new (input and search) masks. Archivista 2009/I is fully multi-client enabled, i.e. you can house several very different archives in one database.

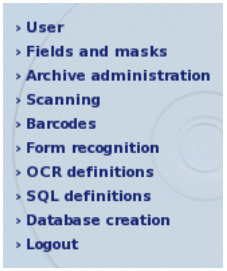
To begin with we describe how to go through the login and logout procedures of the WebClient administration.

13.1 Login

The database that needs setting up is accessed by means of the Internet browser. Host, database, user id and password must be entered. Please note that it is incumbent on you as systems administrator to keep track of all databases and sysadmin passwords.

➲ It may well be that in the framework of your particular installation host and database name need not be entered.

13.2 Logout



To logout you simply click the expression **Logout** at the very bottom of the WebAdmin menu.

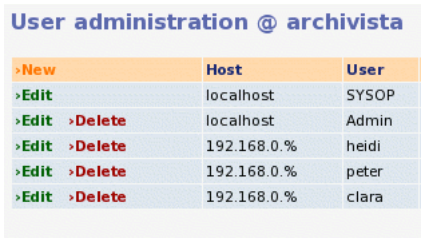
14 User

14.1 Administration (in general)

Tutorial on User Administration (movie 2min) *4.11*

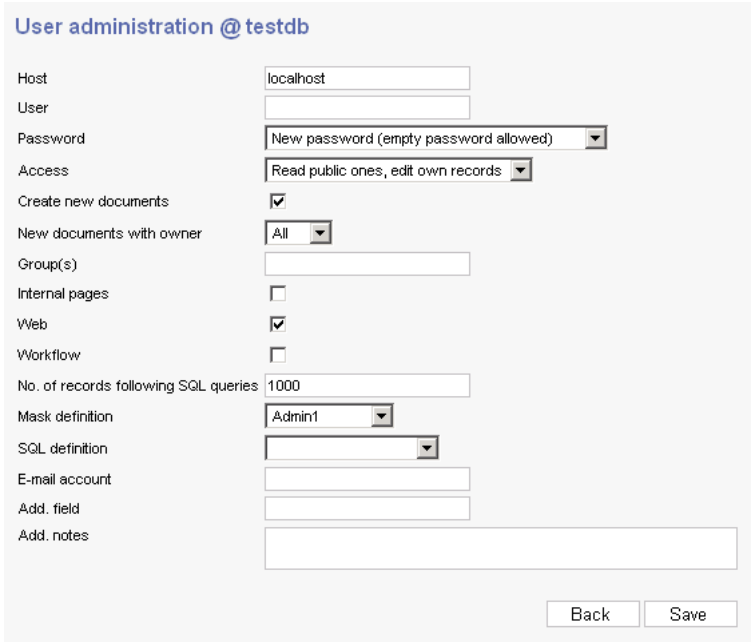
➲ This chapter corresponds structurally to chapter 35.2 under 35. While the settings are made through the web interface here they are made in the desktop application there. If you come across a function or option you do not understand at all, it may help to look up the section in question in 35.

After clicking **User administration** in the menu you see a table of existing users.



(If you have only just started setting up your database, then you see the user 'SYSOP'.) To make changes to existing user definitions click on 'Edit'. To delete users altogether click 'Delete'.

The first cell in the upper left hand corner shows 'New' printed in red. Click 'New' to create a new user. You will see the following dialog box:



The individual settings are described below.

14.1.1 Host

Here you enter the name of the host on which your Archivista database lies. See also 35.2.1.2

14.1.2 User

Here you enter the name of the user you want to create. The user name must be composed of letters but may also contain numbers.

14.1.3 Password

There are three options here:

- Normal login

- New password (empty password not allowed)
- New password (empty password allowed)

When creating a user choose the second or third variant. If you change settings later, then the first variant makes sense. If a user forgets her password, you as SYSOP cannot access it but you set the second option above and the user is forced to enter a new password on next login.

14.1.4 Access

The alternatives are:

- Read public ones, edit none
- Read public ones, edit own records
- Read all, edit own records
- Read all, edit all
- Sysop (all as well as settings)

Each document possesses an owner. Normally the owner is assigned automatically: see 14.1.6. We can imagine many different archives:

1. For instance one with two users and both work exclusively with their own documents.
2. For instance one with two users none of which 'owns' any documents but where there are only public documents.
3. For instance one with three users. Let us imagine a florist, a bookkeeper and a general manager. The florist creates a photo collection for her own purposes, the bookkeeper scans documents posted to an account but puts items of common interest (like the list of public holidays for the new year) to the disposal of everyone and, finally, the general manager compiles records for her own specific uses but is certainly allowed to view all documents.

In the first case both users obtain the following access rights: 'Read public ones, edit own records'. There are no public documents.
In the second case both users obtain the access rights 'Read all, edit all'.
In the third case the users are given the following rights:

- Florist: 'Read public ones, edit own records'
- Bookkeeper: 'Read public ones, edit own records'
- General manager: 'Read all, edit own records'

This way the florist and the bookkeeper see only their respective (and the public) documents. The general manager can view all documents. When the bookkeeper creates a document she would like to make available to the florist, then she sets the owner of that document manually to 'ALL'.

14.1.5 Create new documents

If the user is allowed to create new documents, then this option must be ticked.

14.1.6 New documents with owner

This option is related to 14.1.4. The following alternatives are the most likely choices:

- New documents with 'own' owner
- New documents with owner 'ALL'
- New documents with group owner

See also 29.7.2.

14.1.7 Group(s)

Groups make those documents available to a user that are not really his own but belong to a larger entity. Typical groups are for example the accounts payable department or the production department. Let us take the example of Joe Smith who is part of accounts payable. Under 'New documents with owner' Joe Smith could have 'AP' and with access rights 'Read public ones, edit own records' Joe could handle all documents of the accounts payable department as if they were his own. And his colleagues can do likewise.

See also 35.2.1.4

14.1.8 Internal pages

By ticking **'Internal pages'** you determine that for this user the bitmap pages are not loaded from a directory but are loaded straight from the database. This has the positive effect that users no longer require reading rights for the directories in question: administration is simplified and security enhanced.

➲ Please note that users for which this option is ticked cannot create new documents. This option makes most sense for users who view and edit documents by means of the web browser and who do not create documents manually.

14.1.9 Web

The check-box **'Web'** enables user access via WebClient. If it is unticked the user can access the archive database only via the RichClient and not via the Web module. See also 9.

14.1.10 Workflow

By means of the check-box **'Workflow'** you can trigger certain actions when certain events (creation of new document, alteration of document, etc.) take place. Please refer to 8.6.2 for more information.

14.1.11 No of records following SQL queries

See 35.2.1.3.

14.1.12 Mask definition

The option **'Mask definition'** serves to specify the input mask a user sees after starting up Archivista. The definitions for the input masks are made where the fields themselves are created. By being able to select different input masks for different users you can have several archives in one.

14.1.13 SQL Definition

Here normally 'AVSTART' is chosen, i.e. you choose the definition which determines which document selection a user should see after starting up Archivista. The definition itself is made in a separate input form. See chapter 21.

14.1.14 E-mail account, additional field and additional notes

You can collect one e-mail address per user. The e-mail address may be used by third-party modules (e.g. workflow module) when

certain events trigger user notification. 'Additional field' means that you can allocate an additional identification number to a user, if need be, and in the memo field 'Additional notes' you can enter any text you like.

14.1.15 Additional field

For those users who need to be able to make printouts the administrator can add a printer in the additional field. It must be pointed out that up to now only postscript printers are supported. The printer must be set up correctly as CUPS printer on the ArchivistaBox.

14.2 User administration (external)

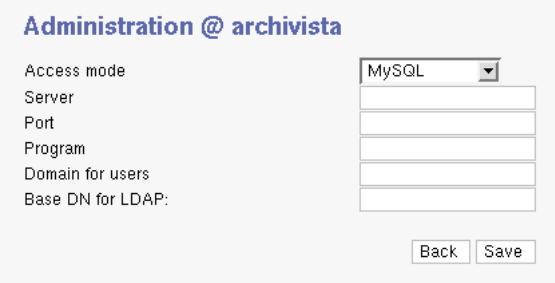
There is a distinction between the external user administration and the internal user administration with MySQL. If you log in to your ArchivistaBox, for example, you directly access the MySQL database i.e. you work internally.

In all those cases where a central user administration tool exists (e.g. LDAP), the ArchivistaBox can be connected to it. At every login the WebClient will check with the computer specified whether the account that is used exists and whether it is valid. Only if this is the case, login can take place.

The way you want to deal with your user administration – and the necessary settings – can be determined under this menu item.

14.2.1 Administration

When you click 'Administration' you will see the following picture:



The first field **'Access with'** gives you the choice of the following options:

- MySQL
- LDAP
- HTTP query

MySQL

If you choose **MySQL**, access will always happen via the ArchivistaBox in question, i.e. no external computer will be used for login.

LDAP

LDAP (Lightweight Directory Access Protocol) is an application protocol that is widely used in computer technics. By means of the LDAP server it is possible to run user administration from one central system. Only one single password is necessary. The web client serves as a data transmission device between central system and the individual database. Hence, LDAP allows you to administrate your users centrally.

HTTP query

The **HTTP** query feature is not yet implemented. However, we are working on it now.

Background information: LDAP allows the query and modification of information of a directory service by means of the TCP/IP network. One important characteristic is its hierarchical setup, which for example MySQL does not use (MySQL possesses an ordinary, tabular structure). The data structure of an LDAP directory is like a hierarchical tree with roots, branches/twigs and leaves. The root (root, suffix) is the topmost data object, below it one finds the higher structures. If you allocate to a user two groups of different hierarchical status (see 14.2.2.1), then the hierarchical status of the user corresponds to the hierarchical status of the one of the two groups which is hierarchically highest.

14.2.1.1 Server

Here you enter the IP address of the server, e.g.: 192.168.0.12

14.2.1.2 Port

If needed, you can enter the interface (port address) here, e.g.: 389

14.2.1.3 Program

If access should take place by means of an HTTP query, you have here the possibility to run a program on a different server. For this purpose enter the name of the program.

14.2.1.4 Basis DN

Here you can specify from which branch of the LDAP tree the search must start. Enter the branch, e.g.: dc=archivista, dc=ch

14.2.1.5 Domain

The mentioning of the domain is necessary in order that the LDAP server may know in which network it is, e.g.: archivista.ch

14.2.1.6 Users small / Groups large

You can use this field to determine that when a user logs in with a mixture of upper and lower case letters, this is always converted to lower case. For example, the user is opened with 'meiere'. If this option is active, 'MeierE' is also accepted. No conversion is carried out if the whole name is entered in upper case.

14.2.1.7 User (Files without owners)

Here, you can explicitly specify a user name so that it is no longer possible to open a file in the ArchivistaBox and all services (scanning, FTP upload and PDF printer) that does not have an owner.

14.2.1.8 Setting a user for uploaded files

If this option is activated, the user that is logged in is set as the owner of the file when the file is uploaded.

14.2.2 Template

You will see something like the following picture when you click on 'Template':

Here you can allocate a user to one or several LDAP group(s).

14.2.2.1 Group (external)

Here enter the name of the group to which the user should be allocated. (See also 14.2.1).

14.2.2.2 Profile (internal)

Select the user who is to be allocated to the group specified above by clicking the arrow.

14.2.2.3 Remark

You can add information on the user in question by entering it in this field.

14.2.2.4 Not active

By clicking this check field the settings you have made remain inactive.

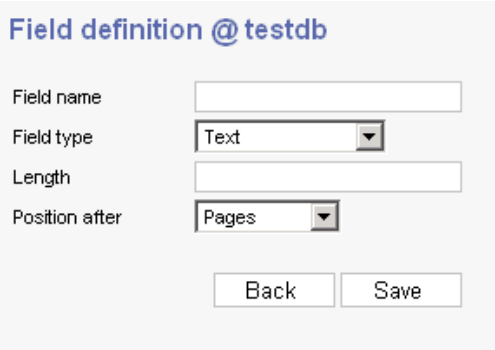
Do not forget to save your changes to the settings!!!

15 Field definition

Tutorial on Field definition (movie 2min) *4.12*

➲ This chapter corresponds structurally to chapter 36.1. While the settings are made through the web interface here they are made in the desktop application there. If you come across a function or option you do not understand at all, it may help to look up the section in question in 36.

After clicking **Fields** in the menu you see a table of existing fields. To make changes to an existing field click **Edit** or **Delete.** The first cell in the upper left hand corner shows 'New' printed in red. Click 'New' to create a new field. You see the following dialog box:



Individual settings are described below.

15.1 Field name

Please note that the name must be unique, must not contain any blank spaces and the first character must not be a number. The name of the field can be changed during the same session but after leaving the form it cannot be changed any more. Should it be absolutely necessary to change it later it can be done from the MySQL console.

15.2 Field type

We differentiate between the following field types:

- Text (1-250 characters)
- Number integer
- Number double (meaning double-precision floating point to express fractions)
- Date in the format yyyy/mm/dd
- YesNo

The field type cannot be modified in later sessions.

15.3 Length

When the field type is 'Text' you may vary the length of the text by entering a value here. Cannot be modified in later sessions.

15.4 Position after

Designates the position of the field within the table with respect to the other fields. Can be altered in later sessions.

16 Mask definition

Tutorial on Mask definiton (movie 2min) *4.13*

➲ This chapter corresponds structurally to chapter 36.2. While the settings are made through the web interface here they are made in the desktop application there. Since with regard to (input and search) masks the graphic user interface is of paramount importance we do **not** recommend to you to take recourse to the RichClient 'companion chapter' in this case. We advise you to adhere strictly to the information below.

After clicking **Mask definition** you see a screen where you can select the mask which you want to manipulate further. The table below shows the fields of the active mask.

16.1 Mask

Here you can specify whether you want to

- select an existing mask and edit it
- delete an existing mask
- rename an existing mask
- create a totally new mask

Before you can make changes to an existing mask you have to select it. To rename a mask you select it first and and then write the new name into the field in front of the button 'Rename'. Then, click 'Rename'.

Please note that you must choose a unique name which does not contain any blank spaces and does not begin with a number.

Whether you create a new mask or make changes to an existing one, there are always the following settings to consider.

16.2 Field name

You can add fields to the mask. Simply pick an existing field from the list.

➲ From reasons of compatibility with the RichAdmin tool the mask definition follows a special kind of logic. In principle, you must add **all** fields to the mask including those that should **not** be displayed. Give these the value '-1' under 'Width'. In this way you ensure that a field does not appear in the mask. See also 16.7.

➲ From the same reasons it follows that a definition field that was allocated to a mask that was allocated to a user cannot be edited by all users.

16.3 Field type

Decide here about add-ons to field definitions. Does a field need to be linked to another field? Do you want to establish a set of fields that are hierarchically dependent on each other?

The following sections describe the different types.

16.3.1 Text code

You can create fields which are linked to other fields. One alternative is the linking of a text field with a text code field. For details see 36.6.4 and 11.3.6.

16.3.2 Normal

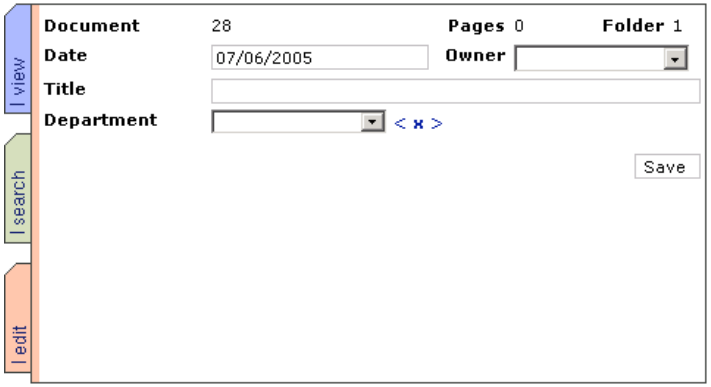
The default value is **Normal**. The field is displayed as a simple entry field.

16.3.3 Definition

The field type **'Definition'** is used whenever we want to have a simple combo box from which to select a predefined value. Such a field is easy to maintain and new values can be added easily too.

The subsequent example elucidates the functionality. Go to 'Field definition' and create a new field named 'Department'. Go to 'Mask definition', select 'Department' from the field list and choose **Definition** as field type. Now save the mask and logout from the WebAdmin tool.

On entering the WebClient you now see the following mask.



Click the field 'Department': as there has not been defined any controlled vocabulary as yet the pull down list is empty.

Click > and a field will open up. Now enter the value 'Purchasing':

Then, click 'Add'. The value 'Purchasing' has been added.

Let us add the values 'Production' and 'Sales' to our field 'Department.' Now you should see a screen roughly like the following:

➲ Please note that when choosing a value from your drop down list it is sufficient to enter the beginning of the word up to the first distinctive character. The value you want is selected and by clicking 'Enter' you transfer it to the field.

16.3.4 1:N

The field type '1:N' enables you to build drop down lists which depend hierarchically on specific entries in other drop down lists, i.e. with this field type you can compile veritable keyword tree structures. Let us make an example. We go to the menu 'Field definition' and create an additional field which is dependent on the field 'Department': let us call it 'Region.' And now we go to the menu 'Mask definition' and choose the essential option: we activate **Field type '1:N'** and we choose 'Department' in the 'Link field'. We have thus specified a 1:n relationship between the fields 'Department' and 'Region.'

Save the changed mask definition, logout from the WebAdmin tool and login to the WebClient. Go to the detail view or input mask in the bottom left corner, click the red 'Edit' tab and under 'Department' choose 'Purchasing.' Now create entries like 'France', 'Italy', 'Germany' and 'Switzerland' in the field 'Region'. Do not forget to save your entries.

Create a new document. Under 'Department' select 'Production' from the combo box. Now go to the field 'Region': you will notice that the controlled vocabulary terms you created earlier ('France', 'Italy', 'Germany') do not appear. They belong hierarchically to the department 'Purchasing'.

For the department 'Production' you could now compile the values 'Switzerland,' 'Germany' and 'Italy,' for example, and select 'Italy' for correct keywording if you like.

➲ 1:N fields can extend over more than two hierarchical tiers. You could, for example, define a 1:n relationship between the field 'Region' and a new field 'Contact.' By doing this you would have created a third tier.

Similarly, you could make several 1:n relationships for one field type 'Definition.' You must simply see to it that the field type 'Definition' is top of the hierarchy.

This is at the same time the most pragmatic approach: move top - down when you want to create hierarchically dependent fields. Always start with the field that is at the top of the hierarchy.

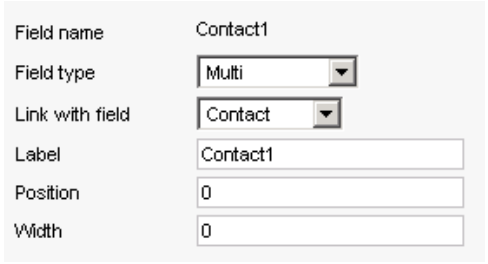
16.3.5 Number code

You can create fields which are linked to other fields. One alternative is the linking of a text field with a number code field. For details see 36.6.4 and 11.3.6.

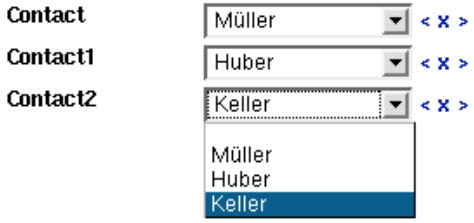
16.3.6 Multi

The idea behind the field type 'Multi' is that with its help you can record several keywords of the same category for a document. Let us assume that you want to assign several contact persons per document. In the following we show you how to do it.

Go to 'Field definition' and create three fields with the names 'Contact', 'Contact1' and 'Contact2'. The field 'Contact' is assigned the field type 'Definition', 'Contact1' and 'Contact2' receive the field type 'Multi' and the entry 'Contact' under 'Link with field'.



The person creating and editing documents can now allocate several contact persons per document. All values predefined under 'Contact' also appear in the combo boxes of 'Contact1' and 'Contact2'. See the example below:



16.4 Link with field

Here you enter the field to which the currently active field refers. For 1:N fields you choose the field with the superordinate concept (in Archivista terminology: 'Definition'). In the case of multi fields, which are in actual fact copies of a main field, the main field is selected.

16.5 Label

Do you want to attach a label to a field? If yes, please enter the label here.

➲ A label makes sense in most cases. With multi fields or hierarchically dependent fields it may be better to omit labels, though.

16.6 Position

Here you can specify at which position the field is to be displayed in the table. Use a number larger than '0'.

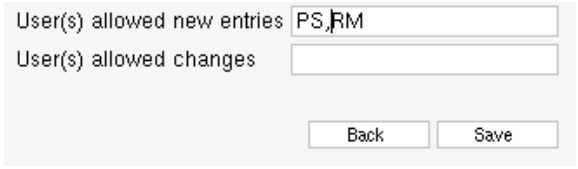
16.7 Width

In the field to the right you specify the column width the field should take in the main view table. One screen pixel corresponds to 15 units, i.e. if you enter 900, the field will be 150 pixels wide.

➲ A value of '-1' means that the field in question is not to be displayed in the table at all.

16.8 User(s) allowed new entries

With the last two options in this dialog box you can lock a field for users that are not listed here.



The users whose names are specified behind **User(s) allowed new entries** can add **controlled vocabulary** to the pulldown list. This option concerns primarily 1:N or multi fields.

➲ Pay attention to separating the user names by commas.

16.9 User(s) allowed changes

Under **User(s) allowed changes** list those users who are allowed to make changes to this particular field. This does not concern changes to controlled vocabulary but solely changes to **selecting from** controlled vocabulary or changes to fields for which entries are not predefined.

➲ Pay attention to separating the user names by commas.

17 Archive administration

Most functions appearing here and described below can also be found in Part IX, in particular under 35.1.4, 35.4 and 35.5. If there is a function you happen not to understand, it may help to refer to the description in Part IX.

17.1 The first options

17.1.1 Suppress field ' Title' in main view

By ticking **Suppress field 'Title' in main view** you decide on a global level (i.e. for all users) that the field 'Title' is not to be displayed. The change can only be seen at the next login.

17.1.2 Width of field 'Title' in table

Under **Width of field 'Title' in table** you can determine the column width that the title takes in the table of the main view.

17.1.3 Hiding extended icons

When you log in to the WebClient, two icons are displayed. One of these is for scanning and the other is for uploading files. If you do not want these icons to be displayed, you have to activate this option.

17.1.4 Field for publishing

The field that you choose under **Field for publishing** is a kind of second owner field with which a document can be allocated to a second owner (user or group user). For many archives this function is irrelevant, however, it can be a very important function indeed. Look at the following instance:

A knowledge archive is continually fed with new documents by two dozen project managers. Each project manager is responsible for the documents she added. At the same time she can make the documents available to an extended circle of managers.

After this option has been activated the input mask will show a second combo box with document owners. This way each primary owner of a document can set a second, subordinate owner. It goes without saying that the main user settings apply also in the extended owner environment. See chapter 35.2.

➲ The field that you specify here must be of the type 'Definition'. See 16.3.3.

17.1.5 Elements in combobox fields (0-x)

The field `Elements in combobox fields (0-x)` gives you the possibility to restrict the number of documents displayed after searches in the WebClient. To define the number of search results displayed enter it in the **combobox field**. Thus, the **combobox field** simplifies and speeds up the search for a specific document in the **WebClient**. (Those interested in getting to know more about searching we refer to Chapter 11.2). The following example may serve to explain how such a restriction works and when it makes sense.

Example: Dr. Lovegood possesses an ArchivistaBox with which he administrates names, addresses and patient histories of his 100.000 patients. Whenever a patient is hospitalized, the already existing data are looked at first. Dr. Lovegood has 20.000 patients named 'Smith' who have undergone an appendectomy. If

he wanted to search for a specific patient without using the **combobox field**, the system would come up with 20.000 hits. For one, this is problematic from the point of view of time and, for another, such a mass of data could cause the system to crash. This is the reason why Dr. Lovegood restricts the number of results to be shown to 100 by entering '100' in the **combobox field**.

➲ If one wanted to see all entries in fields subordinated to the `Definition field` 'Smith' one would have to go to the `Fieldlist`. 8.6.3

➲ To be able to access the field list one needs admin rights. If the **combobox field** is left empty, a search returns all hits.

An entry in this field does not make much sense if Dr. Lovegood has only 5 patients with the name of 'Smith', for example and he restricts the number of results to 20 in the `combobox field` because there were not more than 5 in the first place.

17.1.6 OCR recognition

The **OCR** is a software that serves to recognize text in the case of scanned text. Here you have a choice of four possibilities:

1. Leave the little box unticked if you do not wish any text recognition.

2. Choose **FineReader** if you have purchased an OCR option (until 1/1/2007). This option will yield very good results.

3. If you have not bought an OCR, the OpenSource products **Cuneiform 0.5** and **Tesseract 2.0** offer an alternative.

➲ Both OpenSource products know umlauts, the Tesseract engine can recognize only one language at a time, though (English, German, French, Italian, Spanish and Dutch) That is why with Cuneiform and Tesseract only the first language from the OCR definition (see 20) is taken into account. Ocrad can give very good results if 'Numbers' is selected (makes sense for form recognition).

17.1.7 Create PDF files

The next option is called **Create PDF files**. Ticking this option means that the OCR run does not only make the ASCII text available for fulltext searches but in addition generates searchable PDF files.

17.1.8 Whole document in one PDF file

By turning off the option **'Whole document in one PDF file'** you can effect that upon creation of PDF files each individual page is turned into a PDF file. Normally you should not turn this option off unless you must process extremely large documents with color images. In such a case it could be that the PDF file covering all pages would take more than 128 MByte space and could not be saved in the database because at the moment a maximum of 128 MByte can be stored per database field.

17.1.9 View fulltext field (search form)

When you activate the option **'View fulltext field (search form)'**, then the field for fulltext search is displayed in the WebClient. If the option is not activated, then the field does not appear. Normally, you should leave this option turned on. However, if you do not want to use text recognition and full text search at all (as it can happen for certain types of archives), it makes sense to

deactivate the option as it might bother users to see a field which they cannot use because full text is not available at all.

17.1.10 Show page text in 'View mode'

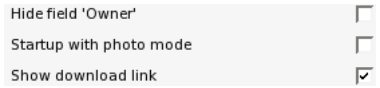
The function **Show page text in 'View mode'** enables you to see nothing but the text of the currently active document in the **WebClient's View mode**.

This function is exceedingly useful if one wants to copy the entire text of the page without any graphics, barcodes or similar.

One should only turn off this function when one wants to put the documents on the Internet because flaws in text recognition can be annoying.

17.2 Further options

17.2.1 Do not show owner



Activation of the field **Do not show owner** has the effect that the field **Owner** does no longer appear in the **View** mode of the **WebClient**. The field serves to indicate which user or user group the currently active document belongs to. To which documents a user has access is described in detail in Chapter 14.1.2. Often the document owner is of no relevance. In this case the field need not be displayed.

17.2.2 Startup with photo mode

If you tick **Startup with photo mode**, then the photo mode is active when you login to the database, i.e. in the upper part of the screen you will see the first pages of the documents instead of the table. This view makes most sense for photo collections.

17.2.3 Show download link

When you access the **WebClient** and the field **Show download link** is active, the table shows the field **Download** at the end of each row.

This field offers the possibility to download the currently active document as PDF file. To be able to download a PDF file the function **Create PDF files** need not necessarily have been activated. One can create a PDF file ad hoc. However, if you want PDF creation to happen automatically and constantly in the background and if you want searchable PDFs you must not only tick the box **Create PDF files** but also **Show download link**.

17.2.4 Quality factor JPEG (1-100%)

Color scans are saved in JPEG format. With this option you can determine the quality of the images. 100% is best quality. The smaller the value you enter the higher the compression.

17.2.5 Scaling of preview (0, 10-100%)

Here you can enter the relation of the size of the preview image to the original file. 25 for 25% is a good value to start with. '0'

means that no preview images at all are created. If you work with the RichClient this can make sense. In general, however, it is good to activate 'Preview'; access of the images is much faster even if this option at 25% requires approximately 10 % more disk space.

17.2.6 Conditions for archiving process (SQL)

Here you can specify by means of an SQL fragment that the archiving process should only involve certain documents and not others. As example may serve `Publish␣<>␣''`. This means that only the documents published to other users should be part of the archiving process.

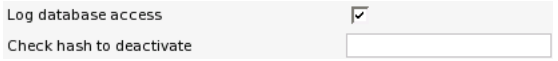
17.2.7 Folder size external storage (MByte)

At regular intervals an Archivista archive undergoes an archiving process. This automated process 'rearranges' and compresses the newly added pages and generally does everything that needs to be done. One important point in this is the preparation of pages for the copying to external media. I.e. in order that you may save your data to a CD or DVD the system prepares a folder of the appropriate size. Under **Size of folder for external storage (MByte)** you specify the maximum size. If you work with CDs, for example, you might enter `600` here.

17.2.8 Max. count files in folder

In addition to indicating the maximum folder size also the maximum number of pages that are supposed to fit in a directory (folder) after the archiving process can be determined.

17.2.9 Log database access



This option enables you to have a log created for all access activities regarding the archive. You find more information on this in 8.6.1.

➲ This option results in the ArchivistaBox working a bit more slowly. Accordingly, turn it on only if you attach importance to having an access log. What is more, you should be aware that the option cannot be turned off so easily once it is turned on, the reason being that otherwise an administrator could hide his or her own accessing the database.

17.2.9.1 Turning off the access log function

Should the access log function accidentally have been turned on or should you wish to deactivate it from other reasons, you can do it by unticking the option 'Log database access' and entering at the same time an 8-digit checking number or check hash. You can take the checking number from the 'access log' table. It corresponds to the first eight characters of the first generated hash (field 'hash').

17.2.9.2 Check integrity of access log

An access log is of little value if we cannot make sure that the log itself cannot be falsified. For this reason keys are generated for each entry in table 'access', keys that later on may serve to confirm the accuracy of the entries. The technology used here was taken from the application 'md5sum'. When so-called 'signing' of the entries takes place, it is always the entries that are signed as well as the hash of the previous entry. This is the reason why it is possible to ascertain later on whether all entries were correctly created. But only if the hash of the first and last (most current) entry are known. Consequently, we recommend that you save the

first hash in a different place after you have turned on the access log function, and to save the most current hash from time to time too.

➲ If you have a good backup concept, it is unnecessary to manually note the hash as these keys are automatically saved in the backups.

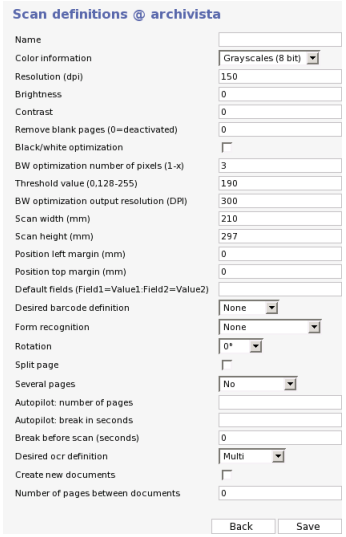
18 Scan definitions

18.1 Introduction

Here, you create various scan definitions that you will repeatedly require during scanning. After you click on the 'Scan definition' menu item, the system displays a list with the existing scan definitions.

You can make changes to scan definitions or delete them by using the 'Edit' and 'Delete' functions which are highlighted in green and red respectively. In the first box on the upper-left hand side of the table, the word 'New' is displayed in red. Click on 'New' to create a new scan definition. The following screen is then displayed:

18.2 General settings



18.3 Settings

18.3.1 Name

Enter the name of the scan definition here, e.g. 'A5 color landscape'.

18.3.2 Color information

Here you specify the color depth, i.e. you decide if the page is to be scanned in black and white, grayscales or color.

18.3.3 Resolution (dpi)

Under **Resolution** you indicate the scanning resolution. It goes without saying that your scanner must support it. Normally you use 300dpi for black and white pages and 150dpi for grayscale and color pages.

18.3.4 Brightness

Normally the value here is 0. If you have to scan very bright or very dark pages you can increase or decrease brightness.

18.3.5 Contrast

Also for contrast the standard value is 0. If you want to increase contrast, try entering 10, 30 or 50. If you want to decrease it, you can enter a minus value.

18.3.6 Remove blank pages (0=deactivated)

This function makes most sense whenever you want to scan your papers recto and verso although only a third (for example) of the stack has any information on the reverse side. Enter the conditions for the removal of blank pages in the following way:

The first is a threshold value for the amount of black on the page expressed in one-tenth of a percent. '20', for example, means that any page whose surface area is less than twenty-tenths of a percent (or 0.02 percent) black will be deleted. (The margin mentioned above is not counted in the calculation of this.)

The second value stands for the margin in pixels that is **not** taken into consideration for the calculation of the proportion of black (as against white) on the page.

0 means that there are no empty pages to delete.

➲ Please note that the two values need to be separated by a full stop. We recommend that you take a figure between 10 and 40 for the first value. The higher the first value the more black a page may contain. The lower the first value, the higher the chance that pages with only a little black information are integrated into the archive. The second value must be '8' or a multiple of it such as '40', '48', etc.

18.3.7 Black/white optimization

When this function is activated scanning takes place in greyscales. As soon as the images arrive at the ArchivistaBox they are turned into black and white by means of an optimization process. I.e. contrasts are reinforced so that the text remains legible even if the original page was of weak print quality.

18.3.8 Threshold (0, 128-255)

With this value you determine how the images are converted back to a black/white raster after the optimization process. The higher the value the higher the proportion of black. If you enter '0', the software will try to remove a colour background.

18.3.9 BW optimization number of pixels (1-x)

The value entered here has an impact on the black/white optimization. It determines the number of points the algorithm reads around the point of origin. The best way to proceed is to test a few typical documents. '3' is a good value to start out from.

18.3.10 BW optimization output resolution (DPI)

Here you determine the resolution the image must have when it is stored in the database. A typical value would be '300'; it is generally recommended for black and white pages.

18.3.11 JPEG compression (factor)

Between scanner and computer (ArchivistaBox) image data are normally transmitted without having been compressed. At 300dpi and color there are ca. 25 MBytes of data per page that must be conveyed. With double sided scanning there are about 50 MBytes per sheet. The USB interfaces used today can in theory handle 60 MBytes (or 480 MBits) per second. In practice the throughput is only about half of that, though, and data transmission may easily take a couple of seconds. There exist a few scanners today which compress the image data while they are still in the scanner and which send out an already compressed data stream. The transferral of data between scanner and computer is much more efficient this way. If the device you use supports such compression you can enter the compression factor here.

➲ Please note that currently only a few Fujitsu scanners offer this option. Use 0 (no compression) or choose a value between 1 (fast transferral with less quality) and 7 (slow transferral, optimum image quality).

18.3.12 Doublefeed detection

Some scanners offer ultrasonic doublefeed detection. By clicking this check field you can activate it. The ultrasound helps the scanner to notice if there is really only one sheet being pulled through. This helps to avoid paper jams and erroneous feeding of more than one page.

➲ Please note that to date only a few Fujitsu scanners support this option.

18.3.13 Scan width (mm)

The scan width of an A4 page is 210mm.

18.3.14 Scan height (mm)

The scan height of an A4 page is 297mm.

18.3.15 Position left margin (mm)

This value is needed with some scanner models when the width of the page does not tally with the area the scanner covers. For example when you use the ADF to scan pages that are narrower than US letter, it is not sufficient to change the scan width but the left margin must be adjusted too.

18.3.16 Position upper margin (mm)

With some scanner models it may be necessary to adjust the upper margin when scanning special paper sizes. This can be done by means of this function.

18.4 Post editing

18.4.1 Default fields

You can enter default values for specific fields here. These fields are then automatically filled during scanning. The following format must be adhered to here:

```
Field1=Value1:Field2:Value2:script:Field3=Value3
```

18.4.2 Calling a program during scanning

You normally always use `Field name=Value` followed by a colon before the next entry. In addition to fixed fields, you also have the `script` option. Here, the relevant program is called after each page that is scanned.
This program is transferred via the command line to the file number, host (computer), database, user and password of the active file. The program can then register in the system and carry out specific post processing activities. For example:

```
#!/usr/bin/perl

my ($host,$db,$user,$pwd,$lnr) = @_;
use DBI;
my $dns = "DBI:mysql:host=$host;database=$db;";
my $dbh = DBI->connect($dns,$user,$pwd);
if ($dbh) {
  my $sql = "select client from archive where file number=$lno";
  my @row = $dbh->selectrow_array($sql);
  my $sql1 = "";
  if ($row[0] == 300) {
    $sql1 .= "Owner='HHWRW'";
  } elsif ($row[0] == 100) {
    $sql1 .= "Owner='KISRW'";
  }
  if ($sql1 ne "") {
```

```
    $sql1 = "update $db.archive set $sql1 where file number=$lno"
    $dbh->do($sql1);
  }
}
```

The above script adjusts automatically to the 'Owner' field if the 'client' field adopts a value of 100 or 300.
The corresponding script was entered here, for example, with `auto.pl` in the input fields. It must be stored under the following path in the ArchivistaBox:

`/home/data/archivista/cust/autofields/auto.pl`

➲: Logically, the scripts are linked with the barcode recognition in the input fields.

18.4.3 Desired barcode definition

When you enter a barcode definition here that you created earlier (see 19.4.0.3 the barcode recognition in question will be carried out automatically after scanning with this scan definition. This way it is ensured that you can work with more than one barcode definition and the document is 'steered down the right path' from the moment the scan button has been touched.

18.4.4 Rotation

Here you can specify that the page undergoes automatic rotation after scanning. This makes sense for pages with a landscape orientation, for example. Choose from the values 0, 90, 180 and 270 degree.
This function is not implemented at this point in time.

18.4.5 Split page in middle

This function serves its purpose when books are being scanned. For more information see Chapter 45.5.

18.4.6 Several pages

If you want to scan more than one page at one go there are a few options to simplify your task.

- ADF Simplex: simplex scanning with ADF
- ADF Duplex: duplex scanning with ADF
- ADF 270/90: duplex scanning with ADF where recto has to be rotated 270 degrees after scanning and verso 90 degrees in order that all pages 'stand on their feet' at the end of the process
- Autopilot: manual scanning with the help of the flatbed scanner; here there are two further settings

18.4.7 Autopilot: number of pages

If you plan to scan a book, for example, you can make the (flatbed) scanner scan at regular intervals while you have nothing else to do than turn the pages. Enter here the number of times you want the scanner to repeat the scanning process.
Diese Funktion ist zurzeit nicht implementiert. Der Flachbettmodus kann zwar angesprochen werden, aber es wird nur eine Seite aufs Mal gescannt.

18.4.8 Autopilot: break in seconds

Here you indicate how long the scanner should wait before doing the next scan.
This function is not implemented at this point in time.

18.4.9 Break before scan (seconds)

If the scanner is located not on your desk but further away you can enter delay time here.
This function is not implemented at this point in time.

18.4.10 Desired OCR definition

Choose here the OCR definition which should be applied to the pages scanned with this scan definition.

18.4.11 Create new documents / Number of pages between documents

These two points belong together. Here you specify that a new document must be created at regular intervals and after how many pages this must be done.

Example Let us assume that you scan invoices which consist of one page as a rule. By activating 'Create new documents' and entering '1' after 'Number of pages between documents' you can build a scan definition that allows you to put a whole stack of paper into the scanner: after each page a new document is created automatically. The scanned 'stream' of pages is thus separated into individual documents without the help of barcode technology. If you want to scan invoices that come as two pages, you can create a second scan definition (and enter '2' beside 'Number of pages between documents).

19 Barcodes (Archivista Box)

19.1 General

19.1.1 Barcode processing with Archivista

The purpose of the Archivista barcode solution is an automated handling of (scanned) documents that are marked or labelled with barcodes. There are three main areas to consider: barcode entry and labeling, barcode recognition, and analysis of the recognized barcodes. After a short introductory part these three areas form the core of this chapter.

19.1.2 Prerequisites

The barcode recognition of Archivista 2009/I (or higher) is based on the barcode engine of Softek Software.

19.1.3 Previous versions

Version 1.0 was the first published version of Archivista Barcode-Print. **Version 2.0** offers simplified entry of barcode definitions and verification mechanisms.

19.2 Barcode technology

19.2.1 Barcode processing: why?

Scanning is only the first step when keeping an archive. Fast retrieval of your data when you need them is equally important. Archivista 2009/I offers you both full text retrieval and the entry of keywords. Full text retrieval is especially useful for archives which lend themselves well to random searches by individual words or word combinations. Working with keywords makes sense in all cases where the filing system in the computer adheres rather closely to a classical paper filing system. Barcode processing goes a step further by efficiently handling large masses of stereotyped documents.

To put it simply, we mark out each document with some unequivocal piece of information according to which a later search may be executed. We achieve this by adding a reference number to our document before we print it or by sticking a barcode label (with reference number) on the document to be filed. As a consequence each document carries a barcode and reference number when it is scanned.

Scanned pages are subsequently worked on by a barcode recognition software and from now on they are at our disposal for searches by reference number or further data analyses. An enormous time saving potential is realized if documents are filed and searched with the help of barcode technology.

19.2.2 What is a barcode?

Barcodes consist of black bars and white spaces. The sequence of these bars and spaces differing in width serves to encode information which can be read (by a scanner) and processed (by a computer) in an automated fashion. The relatively simple encoding allows barcode reading devices to recognize the information with 100% accuracy.

19.2.3 Which barcode types must be distinguished?

There is quite a variety of formats in use. Common to all of them is their (for lay persons undistinguishable) appearance of black bars

and white spaces of different widths. We differentiate between two types of barcodes: two level codes and multilevel codes.

Two level barcodes have two different widths: narrow and wide. Interleaved 2 of 5, MSI Code, Codabar, Code 39 and Extended Code 39 are examples of two level codes.

Multilevel barcodes have more than two variants of width, they may have bars and spaces that are very narrow, narrow, normal, wide and very wide, for example. Among the multilevel barcodes we count Code 11, UPC/EAN and kin, Code 128 and Code 93.

19.2.4 Which barcodes are supported by Archivista?

A barcode solution for the purposes of document management must be able to allocate barcodes to documents and vice versa. To follow this purpose Archivista offers the most widely known barcodes but does not support each and every barcode format. The Archivista Barcode Option works with the barcode formats Code25, Code39, Code128, EAN8 and EAN13. The auxiliary program BarcodePrint supports Code25, Code39 and Code 128. Out of this reason these three formats are presented below:

Interleaved 2 of 5: numeric only barcode, very efficient. Normally you use this format because it can represent high numbers in limited space.

Code 39: alphanumeric barcode. Its ability to represent not only figures but text as well makes it more space consuming than Interleaved 2 of 5.

Code 128: alphanumeric barcode. A space saving multilevel format which yields relatively narrow, unobtrusive barcodes even if figures **and** text are encoded.

➲ A detailed description of the above formats you find on the web page of Hewlett Packard at www.hp.com or on other websites.

19.3 Barcode entry

19.3.1 Barcodes exist on documents to be scanned

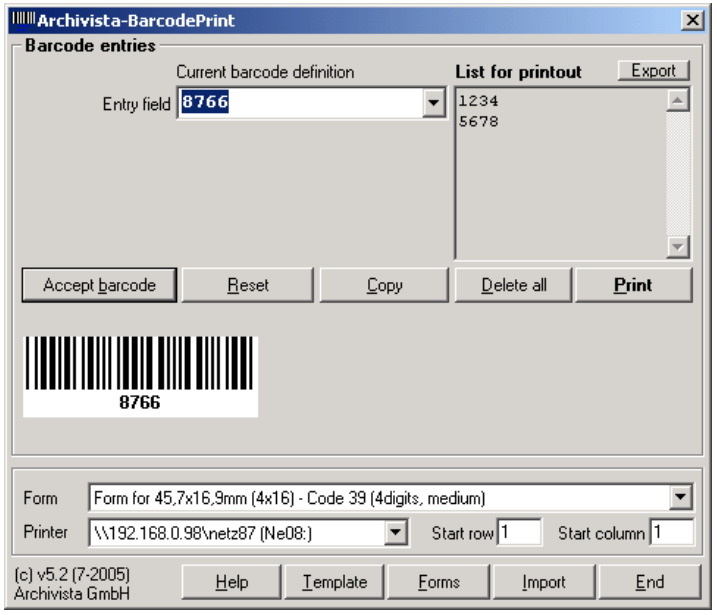
If the barcodes are affixed to the documents during the printing process, you can scan and analyze the documents with the help of Archivista 2009/I and its barcode application without further ado.

19.3.2 Working with Archivista-BarcodePrint

Tutorial on Archivista-BarcodePrint (movie 2min) *5.6.3*

If the documents to be filed do not yet carry a barcode, the barcode information must be added (normally by sticking on a label). This must take place before scanning, otherwise no barcode processing is possible!

To do this, we recommend the use of the sub-application or auxiliary program BarcodePrint of which we show the main mask below:



19.3.2.1 Barcode entry

Working with Archivista-BarcodePrint is simple. After starting the program you get to the entry mask. Now enter the barcode information in the **'Entry field'** and the corresponding barcode becomes visible below. By pressing the Return key the entered reference number is copied to the list for printout on the right. The field 'text for barcode' is ready for the next reference number which again you copy to the list on the right by hitting the Return key. When all reference numbers are entered you can print the barcodes by clicking the field 'Print'.

➲ Barcodes that have been entered previously may be called up once again by clicking the combo entry field (black triangle on the right).

19.3.2.2 Editing and deleting individual barcodes

Entered barcodes are moved to the 'list for printout'. There you can edit and delete them if you wish. To do this, click the area **'List for printout'** and correct or delete the entry in question.

➲ After creating a barcode definition you can click on **'List for printout'.** A dialog box appears asking you **'How many barcodes would you like to create?'.** Enter a number and the corresponding number of barcodes will be generated: the last digit will increment by one for each barcode.

19.3.2.3 Copying a barcode to the clipboard

It is possible to copy the barcode to another Windows application. Simply click the field 'Copy' after entering the barcode information and paste it to any Windows application which is able to handle bitmap files.

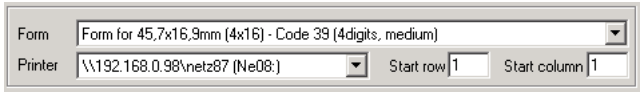
19.3.2.4 Printing of entered barcodes

To print all barcodes appearing in the list click 'Print'.

➲ Printing takes place without further confirmation!

19.3.2.5 Printing options

If Archivista-BarcodePrint has once been adjusted to your needs, you normally do not have to change any settings. However, once in a while you might need the printing options to achieve the desired printing result.



First, choose the desired form. The form active during the last session normally appears as standard form at the beginning of the next session. You create new forms by clicking the button 'Forms'. You can also edit forms, for example, to adjust the page layout.

Choose the right printer under **'Printer'**. After starting up Archivista-BarcodePrint the standard printer appears.
The options **'Start row'** and **'Start column'** are needed to help you saving label sheets. Printing starts only in the mentioned row and column.

19.3.2.6 Using label sheets

To achieve best results barcodes should be printed with a laser printer, which gives the bars a sharp edge. If the barcodes are printed in a large format an ink jet printer may also serve the purpose.
Ordinarily you use self-adhesive label sheets. By default Archivista-BarcodePrint works with the label sheets from the Zweckform company: **Zweckform No 3667, 4732 and 6071.** The standard forms were created for these label sheets.

19.3.2.7 Deleting all added barcodes

After printing the barcodes the individual barcode entries are still in the printing list. You can print them again or edit them if you like.
If you do not need them anymore you delete them by clicking the field 'Delete all'. All entries in the printing list are erased.

19.3.2.8 Importing existing barcode information

If barcode information exists already in electronic form you can import it via an ANSII file directly into Archivista-BarcodePrint. Use the button **'Import'.**

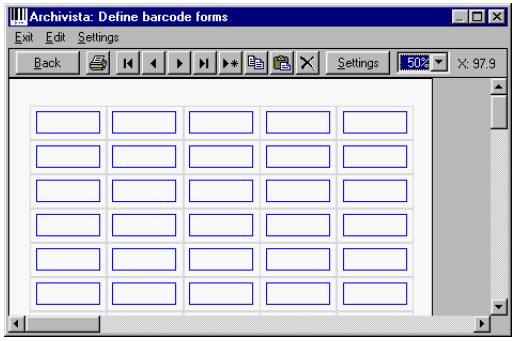
19.3.3 Archivista barcode forms

If you cannot print the desired barcodes by means of the existing forms, you can edit these forms or create entirely new ones.

➲ The Archivista barcode form generator is very similar to the report generator of Archivista 2009/I. Please also refer to 'Define print reports' in the Archivista manual (chapter 37).

19.3.3.1 Form generator

¿From the main mask you get to the form generator by clicking the button **'Forms'.** You will see the following:

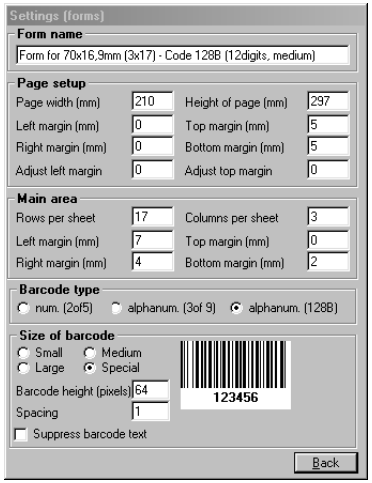


All functions with regard to adding and deleting individual forms you find in menu 'Edit'.

➲ The chosen form is shown on screen. Each little blue box represents a barcode field.

19.3.3.2 Making form definitions

You can define how a form looks by going to the **form generator** in menu 'Settings' / 'Page settings' or directly by clicking the button **'Settings'** in the icon bar. Now you can enter the settings according to your needs.



Under **'Form name'** you give a name to your form. This information appears later in the main mask where you can choose among the different forms defined by you.

Under **'Page setup'** you find all the necessary page information. Please note that the **page size values are not sent to the printer** when printing. The advantage is that you can print with all paper sizes. You must ensure that the values set do not exceed the physical paper size, though. With some printers it may be necessary to move the whole page a little to the left or to the right, or a little up or down. To do this you use **'Adjust left margin'** and **'Adjust upper margin'**

Under **'Main area'** you give the number of rows and columns as well as the margins for the individual barcode cell.

Under **'Barcode type'** you choose the type of the barcode to be printed. Choose between the numeric only barcode 'Interleaved 2 of 5' and the alphanumeric barcodes 'Code 3 of 9' (= 'Code 39') and 'Code128'.

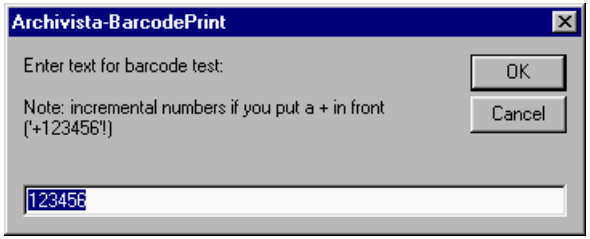
Under **'Size of barcode'** you define how the barcode looks. You can choose between three standards: **'small', 'medium'** and **'large.** If you want to define your own choose **'special'.** The fields **'barcode height'** and **'spacing'** are at your disposal. 'Spacing' influences the gaps between the black bars and the white spaces. You can see the changes you make in the barcode picture on the right of the entry fields. **If you do not want to print the barcode text**, tick the corresponding box.

19.3.3.3 Saving of form settings

Changes to forms are registered in the memory after you confirm your changes before you go to another form or when you leave the form generator. To save the changes and have them ready for further sessions you must end the program. Only at that point the new form formats are written to the file 'AVBCODE.DAT'.

19.3.3.4 Print test page

After you have entered all parameters in the form 'Settings' you should print a test page in the **form generator** to check the looks of it. To do that either click the printer icon or choose the option 'Print test page' in menu 'Exit'.



Enter the text of one barcode as example and click 'OK'. After that a whole page with the same barcode is printed to give you the possibility to check margins, readability by scanning it, etc.

➲ To achieve incremental numbering you put a plus in front of the barcode information (e.g. '+123456').

Problem solving: if the printout does not give the desired result, you must alter the form settings accordingly. Problems occur often

because the margins are set below the minimal printer margins. In such a case you simply increase the margin in question.

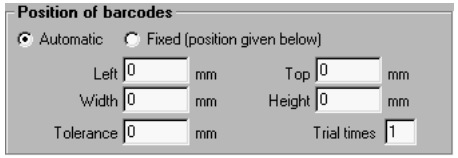
19.3.4 Sticking barcode labels on documents

After printing the label sheets with Archivista-BarcodePrint you stick the labels onto the documents. There are two ways to do that:

19.3.4.1 Labels in variable positions

Normally you work with this option. In principle you can affix the label to the document anywhere you like. However, you should take care to place it in a way that there are a few millimeters of space between label and surrounding text.

➲ If you work with this option you must tick it in 'Options for barcode recognition' ('Database', 'Parameters' and 'OCR settings'). See illustration:



Problem solving when barcode is not recognized: with the 'automatic' recognition of the barcode, i.e. when the recognition software locates the barcode itself, problems may occur when the barcode is not long enough. If you work with format 'Interleaved 2 of 5' and 6 digit barcodes this may be the case. You best change to the longer format 'Code 3 of 9' / 'Code 39' to avoid this problem.

19.3.4.2 Barcode labels in the same position

It may make sense to position the barcode label on all documents in the same place. In particular if there is other barcode information on the document or if the barcode is so complex that 'automatic' recognition presents problems.

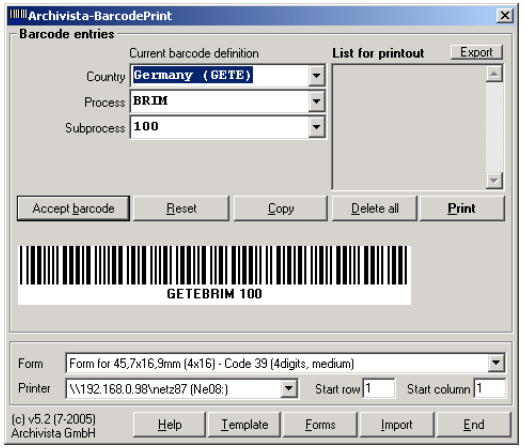
➲ If you work with this option you must tick it in 'Options for barcode recognition' ('Database', 'Parameters', 'OCR settings'). Choose 'fixed' and enter the details of the desired barcode position.

19.3.5 Define your own BarcodePrint forms

Tutorial on Define your BarcodePrint Forms (movie 2min) *5.6.2*

It is not difficult to tailor the BarcodePrint form to your needs. However, we recommend that you study this subchapter at some calmer moment since the possibilities are rather comprehensive. Investing some time here may save you enormous amounts of time later on.

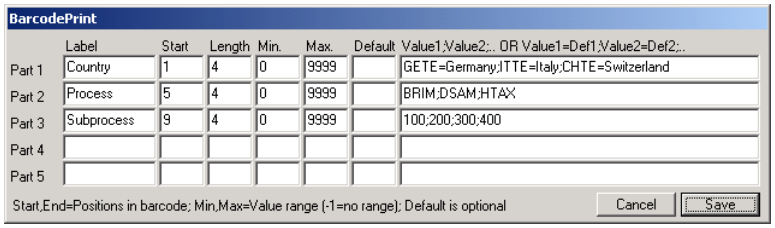
When you start the Barcode program after having installed Archivista 2009/I you will find exactly one field into which a barcode may be entered. The illustration below shows you an example of how the BarcodePrint form may be adjusted:



You see three fields which together generate one barcode. The parameter file **'avbcode2.dat'** serves to check whether the barcode is correct. This parameter file lies in the program directory of Archivista BarcodePrint.

19.3.5.1 Working with templates

Thanks to the button Templates you can easily adjust the barcode entry form.



Each 'analysis unit' of the barcode takes one row. Take care to separate the values in the last field by means of semicolons.

Alternatively, you can adjust the entry form directly in the parameter file 'avbcode2.dat'.

Within a parameter file a maximum of five entries are possible because there are only five fields. The lines 2 to 6 of the parameter file are decisive. Each row is built along the same lines:

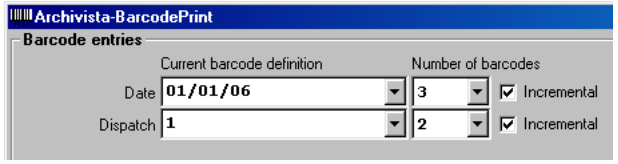
- **Label:** text that is supposed to appear in BarcodePrint form, e.g. Country
- **Starting position:** specifies what position this entry takes in the barcode; the first entry starts at '1'
- **Length:** the maximum length of the entry; a '3' means that this part of the barcode consists of three characters
- **Min:** minimum value of barcode
- **Max:** maximum value of barcode
- **Default:** the standard value (must be between 'Min' and 'Max')
- **Value1;Value2;... (optional):** here you can specify up to 100 precise entries that should be accepted. The format of this list is the following: 'Value=Description;'. For 'value' you enter the figures or letters that are part of the barcode, for 'description' the name you want that analysis unit to stand for.

➲ With the help of the parameter file 'avbcode2.dat' you may create barcodes for the formats 'Code25', 'Code39' and 'Code128'. Alphanumeric characters (i.e. with letters) are supported by 'Code39' and 'Code128'! Before you can create a new template you have to make sure that in the barcode **form** (press button 'Form') the right barcode type is specified.

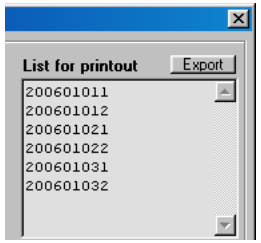
19.3.5.2 Series

In a limited way it is possible to create barcode series.

Let us consider the following example. Twice a day goods are dispatched from your warehouse. The delivery notes are filed in a folder. At the end of the week you would like to scan the documents but without detailed indexing information. Your aim is above all fast indexation. Only the date and a '1' or a '2' for the morning or the evening dispatch, respectively, are required. BarcodePrint enables you to create the necessary barcodes with one click. By entering the following information in the barcode print program a series of barcodes is automatically prepared:

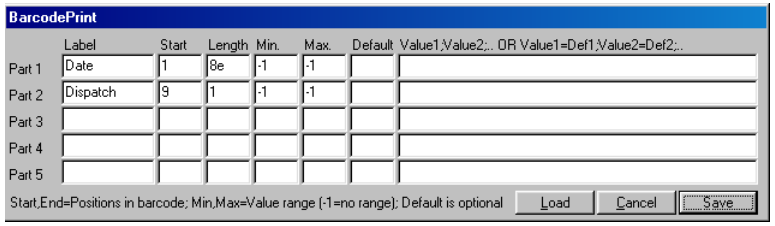


The date can be entered in English format 'mm/dd/yy'. However, the date as it appears in the barcode displays the format 'yyyymmdd'.



With one click barcodes for the whole week have been created: they carry the date followed by '1' or '2'.

The template belonging to the example above determines how the entry mask looks:



The code entered under 'Length', namely '8e', defines that the date will be printed out in 8 digits and that it can be entered according to Anglosaxon custom, e.g. 12/31/06.

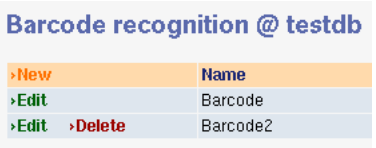
The codes used are

- 6d
- 6e
- 8d
- 8e

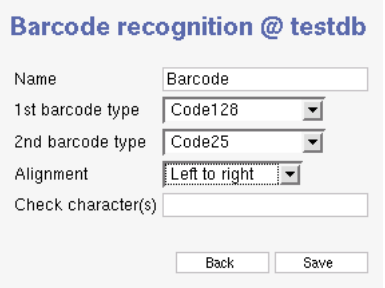
whereby 'd' stands for German entry, 'e' for Anglosaxon entry and '6' or '8' for the number of digits needed for output.

19.4 Barcode recognition

Here you specify different barcode recognition definitions that you think you will need in the future. After clicking the menu item 'Barcode recognition' you first see a list of existing definitions.



To make changes to existing definitions click 'Edit'. To delete definitions click 'Delete'. The first cell in the upper left hand corner shows 'New' printed in red. Click 'New' to create a new barcode recognition definition. You will see the following dialog box:



The individual settings are described below.

19.4.0.3 Name

The system is basically designed for parallel use of different barcode definitions. Thus, you may want to use one barcode definition for vendor invoices and another one for delivery notes signed by your clients, for example. The different definitions are distinguished by names that you enter here.

➲ Please note that as per July 2005 only one barcode definition, i.e. the first one, can be used.

19.4.0.4 1st and 2nd barcode type

'Barcode type' determines the kind of barcode to be recognized. Archivista recognizes the following barcode types: Code39 (with or without test characters), Code25 (with or without test characters), EAN13, Code128 and EAN8. If type is set to 'automatic', every barcode type is recognized.

➲ The barcode recognition can recognize two different barcode types, e.g. Code128 and Code25.

Further accuracy in barcode recognition can be added by setting the **barcode's alignment.** The options are 'left to right', 'bottom to top', 'right to left', 'top to bottom'. When set to 'automatic', any alignment will be recognized.

Check characters give even further control over barcode recognition. However, this function only works if all barcodes to be recognized begin with the same set of ID characters. By entering 'STAT,PERS', for example, you effect that only barcodes beginning with STAT or PERS are recognised as valid. This increases accuracy of barcode processing.

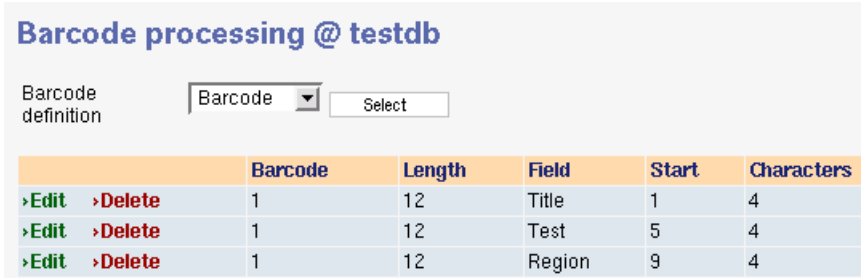
You can use **Stretch vertically** to stretch barcodes of less than one inch in height so that they reach the minimum height.

You use **Recognition after each page** to ensure that barcode recognition occurs even if the same barcode exists on a subsequent page. This guarantees that each page with a barcode is always stored in a separate folder.

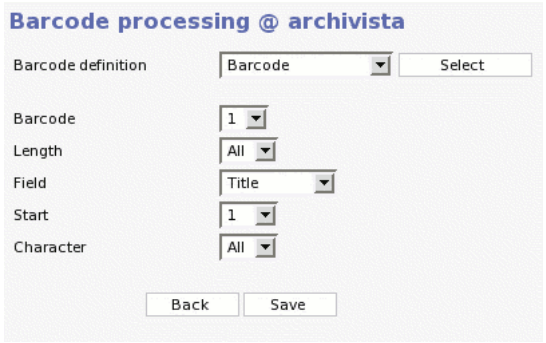
19.5 Barcode processing

Here you specify how the barcodes are to be decoded after text recognition i.e. which part of which barcode is to be allocated to which field.

After clicking the menu item 'Barcode processing' you first see a list. Each line represents an 'analysis unit' or piece of information that is allocated to a specific field. Thus, the information of one barcode may be 'food' for several fields or, indeed, two barcodes may deliver information that fills as many as ten fields.



To delete definitions click 'Delete'. To add definitions or make changes to them click 'Edit'. You will see the following dialog box:



The individual settings are described below.

➲ Please note that what you enter here will later run in the background when you or your colleague scans documents with barcodes on them.

Barcode definition: Choose here to which barcode the following settings apply. (Click the **Select** button.)

Barcode: Here you narrow down further to which of the five barcodes that theoretically can be gathered under the umbrella of a barcode definition the following settings apply.

Length: Enter here the total length of the barcode expressed in number of characters. You can simply choose 'All'. However, the advantage of entering the total length is that the system checks again whether the barcode considered is the right one.

Field: This entry determines the field into which the barcode information should be copied.

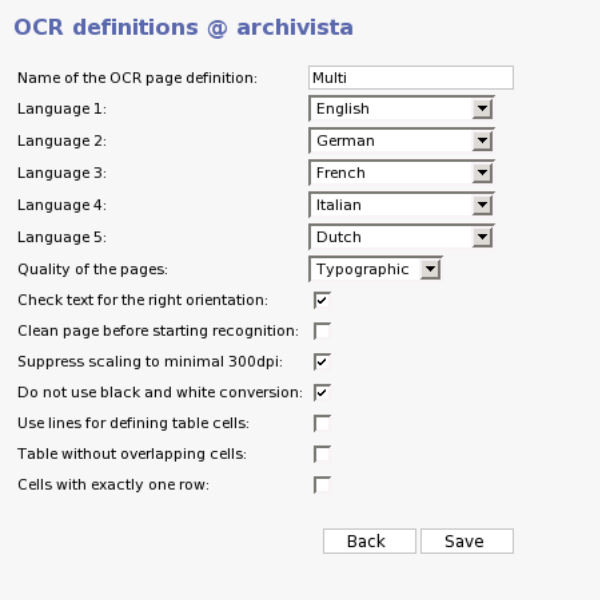
Start: If you want to accommodate the first 'analysis unit' of the barcode, then use '1' for 'Start'. The analysis units that follow have starting values that correspond to their character positions.

Character(s): By selecting 'All' you transfer the entire barcode information to the field specified. By choosing a value between '1' and '20' you determine the number of characters entered in the field in question.

Please do not forget to save your settings.

➲ Beside text and numerical fields you can also address date fields. However you must use a 6-digit or 8-digit format for the date: it must be all numeric without separating characters of any sort. For example, '040130' would work and '20040130', too, but '30.01.2004' would not.

20 OCR definitions



The form **'OCR settings'** enables you to get the most out of the internal text recognition module. You may define up to 20 sets of definitions which may be allocated individually. The buttons at the bottom of the form help you to jump from one definition to the next or to delete a OCR definition.

20.1 Name of OCR page definition

Each OCR definition must wear a name which you specify in line 'Name of OCR page definition'. This name will appear in submenu 'OCR settings current page' which may be used to assign a specific definition to a specific page, document or selection of documents.

20.2 Languages of definition

Each definition may contain up to five languages. There is a minimum of one language. The specification of several languages helps to maximize the number of recognized symbols. If, for example, 'German' and 'French' are entered, the umlauts of both languages are valid typefaces but an Italian umlaut will not be rendered as such.

20.3 Text quality of pages

Normally, you will use the option **'Typographical'**. If you have texts which stem from a matrix printer or a typewriter the other two options help you to achieve better quality.

20.4 Further preparatory settings

'Check text orientation' means that pages that were scanned in the wrong direction are rotated so that OCR is possible. 'Clean page (e.g. delete background pattern)' deletes screen dots that otherwise would get in the way of text recognition. The option 'Suppress minimal scaling to 300dpi' prevents pages that exist in a weaker resolution from being extrapolated to a minimum of 300dpi. 'Process without black and white conversion' means that pages in gray scales or color are not converted but recognized as such.

20.5 Options for table recognition

If you have pages or documents which contain tables these options help you to make the most of them. They are especially valuable

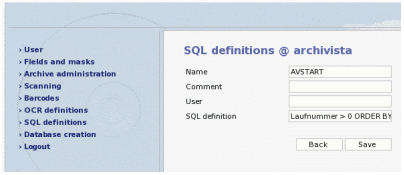
in cases where you want to export the information contained in the tables to another Windows application.

The option **'Define cells of table according to lines'** 'tells' the OCR software that only cells with borders should be recognized. If the internal OCR does not recognize a table properly just take the original page, draw lines where you want to have them, scan the page again and have it recognized once more. In 99.9 % of all cases this will do the trick.

The option **'Tables does not contain overlapping cells'** ensures that no cell extends over another. If your table shows titles that go over more than one column this option should be turned off.

The option **'Cells contain exactly one row'** means that all cells are recognized as having one line although borders or spacing may be missing.

21 SQL definitions



The SQL definitions described here serve the purpose of giving specific users specific selections at start-up. The name of the SQL definition must be AVSTART in all cases.

The individual fragments must be separated by a space each and are to be used in the sequential order described below.

```
SELECT * FROM Table WHERE (Conditions)
ORDER BY (Sort)
```

➲ The underlined part of the SQL fragment must not be entered. It will be added automatically when a selection is made via menu 'Selection'.

(Conditions): The conditions serve to restrict the selection of records. You can use the ('%') sign as wildcard symbol, however, you must use it according to SQL standard. Leaving away the condition 'WHERE' would result in all records being selected.

(Sort): After the fragment 'ORDER BY' the field names follow according to which the sorting takes place. You can use several fields. The abbreviations 'ASC' and 'DESC' result in ascending and descending sorting, respectively. 'ASC' is optional, i.e. when nothing is mentioned an ascending sorting is effected. If you omit the condition 'ORDER BY', no sorting will take place.

Example:

```
Laufnummer > 0 ORDER BY Datum DESC, Laufnummer DESC
```

22 Form Recognition (ArchivistaBox)

22.1 Introduction

With form recognition it is possible to automatically process scanned pages that have the same structure or arrangement of items on the page.

Let us assume that we would like to archive suppliers' invoices. The look and structure of the invoices is the same in all cases, i.e. the invoice date and the invoice number are to be found in fixed positions. We would like to extract both pieces of information from the invoices of the same type (forms) and copy them into the pertinent fields of our archive and we would like to split the pages automatically into documents.

This means that we scan all invoices as one batch, then the form recognition module splits them into individual documents, assigns automatically invoice numbers and dates, and files them in the archive.

In order that ArchivistaBox can do this, we must enter forms in `WebAdmin`. As a second step these forms must be activated in scan definitions (see 18). The actual start of the form recognition process coincides with the triggering of the scanning process in the WebClient (see 12.2).

The form recognition module allows us to manage several forms. These are defined in `WebAdmin` and menu item `Form recognition`. One can add individual forms as so called masks. One mask corresponds to one form definition. Within a definition there are objects telling us where on the form exactly a piece of information can be found (e.g. invoice number), to which type it belongs (e.g. numerical field) and in which field of the archive it should be stored.

Thanks to good text recognition software the extraction of the information works well, however, we encounter the problem that scanning invariably involves certain inaccuracies. The invoice is not scanned straight, there is paper skew or, the page was printed awry in the first instance.

For all these cases there is the `logo recognition`. Thanks to it the first step is a logo search on the page and a subsequent positioning of the objects in accordance with the logo. This enables the recognition process to extract the needed information in a very exact manner no matter how much paper skew there was during scanning.

➲ Whenever possible the logo recognition should be used together with the form recognition. If this is not a possibility because the pages do not contain a logo, for example, one must add tolerance to the areas to be recognized (ca. 3 millimeters).

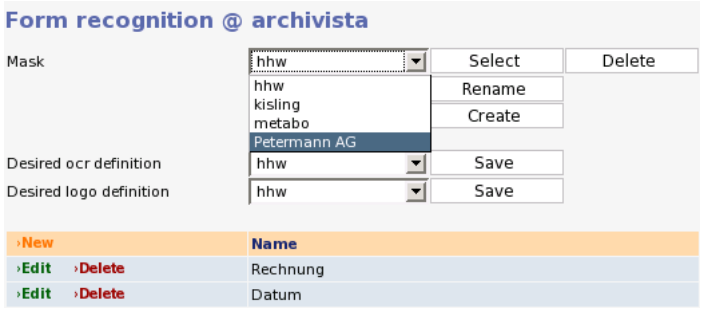
In the following the functions in connection with adding forms are described. Each mask definition corresponds to a form definition.

22.2 Managing forms

Form recognition allows you to manage several form definitions.

22.2.1 Selecting a definition

To activate a definition click on the definition name in question in the drop down list. Then click `Select`.



22.2.2 Deleting a definition

The activated definition can be deleted by clicking `Delete`.

22.2.3 Renaming a definition

The activated definition can be renamed. This is effected by clicking `Rename`.

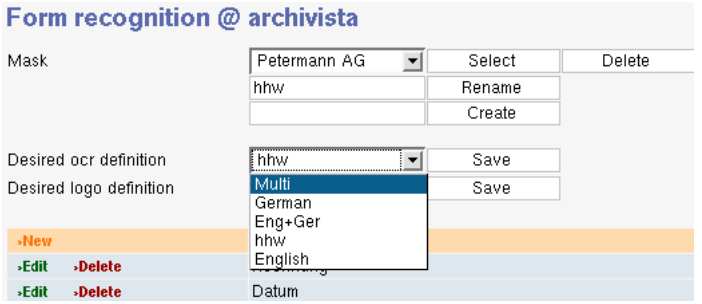
22.2.4 Creating a definition

If you want to add a new mask, enter its name in the field preceding the button `Create` and press `Create`.

22.2.5 Allocating OCR

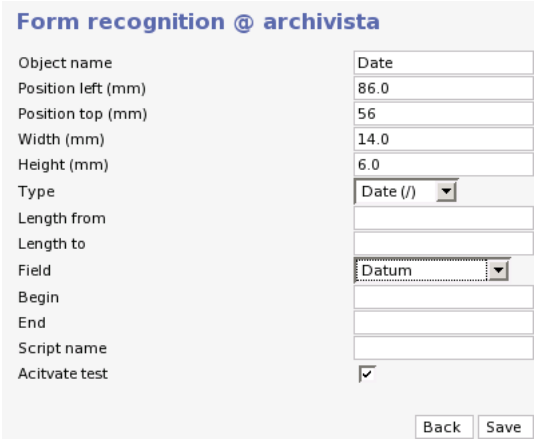
For the extraction of the objects each form definition is in need of a text recognition software. That is the reason why each form recognition must be allocated an OCR definition (see 20).

➲ If there are only numbers to be extracted (e.g. invoice number and date), it makes sense that for the OCR definition only 'English' is chosen or, if available, 'numbers'.



22.3 Editing objects of currently active definition

By clicking on `New` a new object can be created in the active definition.



The following properties can be determined:

Object name

Each object carries a name. This serves for organisational purposes but it does not influence the recognition.

Position

By means of the elements `Position left` and `Position top` the distance from the left and the upper page margins may be defined. The upper left corner constitutes the point of origin. Thereby the page length corresponds to the y-axis and the page width to the x-axis.

➲ **Example:** The invoice number can be found 9cm from the left and 6cm from above. If you work with logo recognition you enter 89mm for 'Position left' and 59mm for 'Position top'. If you work without logo recognition, you must calculate approximately three millimeters of additional leeway, i.e. enter 86mm and 56mm, respectively.

Width/Height

The measures of the object can be determined with the help of the elements `Width(mm)` and `Height(mm)`.

➲ **Example:** The invoice number is exactly 1.2cm wide and 0.4cm high. In analogy to the example regarding position above one would enter 14mm for width and 6mm for height in the case of logo recognition, or 20mm for width and 12mm for height if one works without logo recognition.

Type

With this field you can determine of which type the piece of information is that you want to capture. There are the following choices:

- Text: entire text is accepted.
- Numbers: The first number in the recognized text is considered valid.
- Date (.): The first date in format d(d).m(m).yy(yy) is extracted. Examples: 13.4.07 or 20.12.2006
- Date (/): The first date in format m(m)/d(d)/yy(yy) is extracted. Examples: 04/13/07 or 12/20/2006

Length

The elements `Length from` and `Length to` enable you to restrict the length of the object. If, for example, only six digit numbers are to be recognized, you enter '6' in both cases. If six to eight digit numbers are to be accepted you enter '6' in field `Length from` and '8' in field `Length to`.

➲ If length is unknown, leave the elements empty.

Field

Element `Field` determines to which field in the archive the recognized text of the object is copied.

Begin and end

With these two options you can define which parts of the recognized text are to be copied to the previously specified field.

➲ **Example:** We have the number '12220001' and would like to extract only the first four digits. In this case enter '1' for element `Begin` and '4' for element `End`.

Script name

This option is used to process (further improve) the recognized text. By entering a script name here you effect that during the form recognition process the mentioned script is run. As first and only parameter the script takes the file name of the recognized text. The script must open this file, process the text and display it on the console. The form recognition takes the script output and files the value in the field in question in the archive.

➲ **Example:** From a date that was recognized only the year should be extracted. The Perl script below does just that job:

```
#!/usr/bin/perl
use strict;
my $file = shift; # get the file name
open(FIN,$file); # read the file
binmode(FIN);
my @lines = <FIN>;
close(FIN);
my $txt = join("",@lines); # get the whole text in one variable
$txt =~ s/\r/ /g; # replace all return with space
$txt =~ s/\n/ /g; # " all newlines
$txt =~ s/\t/ /g; # " all tabs
$txt =~ s/\s\./ /g; # " space and point goes to point
$txt =~ s/\.\s/ /g; # " point and space goes to point
$txt =~ s/\s{2,2}/ /g; # replace two 2 spaces with 1
$txt =~ /^(.*)([0-9]{2,2})(\.)([0-9]{2,2})(\.)([0-9]{4,4})(.*)$/;
if ($2 ne "" && $4 ne "" && $6 ne "") {
  $txt="$6"; # if we got a day,month and year, give back the year
} else {
  $txt=""; # don't give back anything
}
print $txt; # print it out (give it back to the form recognition)
```

In order that the script may be run it must be stored in the path mentioned below.

`/home/data/archivista/cust/formrec`

To copy a script into this folder you can proceed in the same manner as when you prepare a logo for form recognition. For more information see 22.4.2.

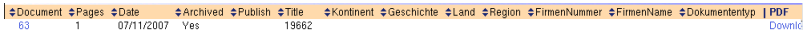
Activate test

With this option you can turn on a test. If it is ticked, the form recognition module creates a PDF file for each page showing you how the object in question is handed over to the OCR software. Positioning mistakes are easily traced this way.

It may happen that when the area in which the object to be recognized is is specified too small or too big, an object and with it the entire form is not properly recognized.

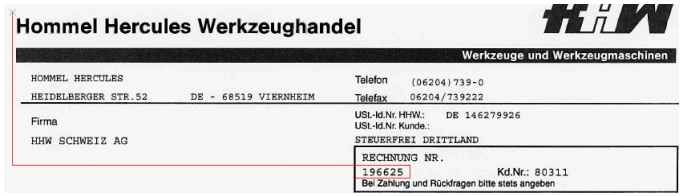
To execute a test turn on this option. Note that only one object can be tested at a time. Then, scan a few pages with form recognition turned on. Wait until the pages have been processed. After that you can open the generated PDF file in the `WebClient` and check whether the object has been positioned correctly.

➲ **Example:** In the following form the invoice number filed in field `Title` is wrong. It should be 196625 but we see in the WebClient that only the number 19662 appears.



Document	Pages	Date	Archived	Publish	Title	Kontinent	Geschichte	Land	Region	FirmenNummer	FirmenName	Dokumententyp	PDF
63	1	07/11/2007	Yes		19662								Downlo

When you look at the PDF file (you can do this by clicking the field `Download` at the end of each line), you see by the fine gray line (dyed red for illustration purposes) what the reason may be.



When you look at the frame around the invoice number, you find out that it is situated a bit too much to the left and that it includes the line which it adjoins. Such a line may prove to be a stumbling block for the text recognition software.

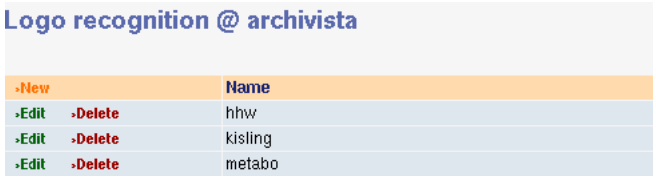
22.4 Logo recognition

Logo recognition is a subcategory of form recognition. When scanning it may easily happen that there is paper skew. This in turn may cause troubles for the form recognition as it is dependent on the exact position of the objects to be recognized.

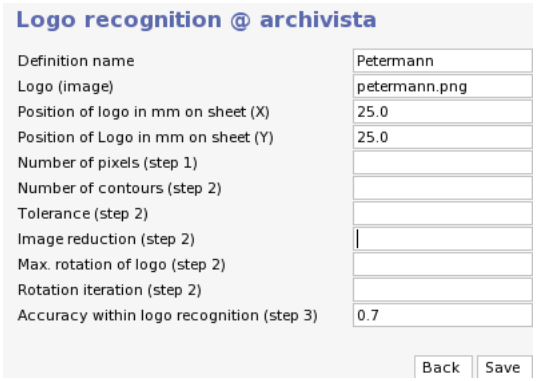
Logo recognition means that the logo previously stored in the system is recognized on the page. On the basis of the logo's position the system is enabled to decide whether there was paper skew during the scanning process and to correct this, if necessary, in order that the form may be recognized without any problems.

22.4.1 Creating a logo recognition definition

To create a new logo recognition definition you proceed as follows: in menu `Form recognition` in `WebAdmin` you find the item `Logo recognition`. When you click it you see the following:



By clicking **Edit** you can edit existing logo recognition definitions and by clicking **Delete** you can erase existing definitions. To create a new definition click **New**. You will see the following mask on your screen:



In the following the individual functions are described:

- **Name of definition**

  With this field you give the definition a name. In our example it is Petermann AG.

- **Logo (image)**

  In order that the logo may be recognized it must be stored as image file on ArchivistaBox. Fill in the file name here.

- **Position of logo in mm on sheet (X) and Position of logo in mm on sheet (Y)** Imagine the form as the 4th quadrant of a Cartesian coordinate system in order that the page width constitutes the x-axis and the page height the y-axis. The origin, i.e. the point from where measurements are taken, is in the left upperhand corner where the two axes intersect. The horizontal and vertical distances of the logo from this point on the axis are the values that must be entered in these fields.

- **Number of pixels**

- **Number of contours**

- **Tolerance**

- **Image reduction**
- **Maximum rotation of logo**
- **Rotation iteration**
- **Accuracy within logo recognition**

  As mentioned in 22 an exact definition of the area where the object is to be found is absolutely necessary. In the same vein it is Utopian to assume that it is possible to scan a document without it being slightly skew. This is the reason why we must give the system some tolerance. This field contains the accuracy factor whereby '1' means 100 percent accuracy. In the illustration above you see the factor 0.7 as a workable average.

➲ The options **Number of pixels, Number of contours, Tolerance and Image reduction** cannot be used yet.

After you have filled in the necessary fields you must save your settings.

22.4.2 Creation of a new logo

Firstly, a page containing the logo must be scanned. In order that good results are achieved later during recognition you should take careful note of the following points:

- Scan letterhead page in color at 300dpi.
- Set quality to 90 at 17.2.4.
- Use a high-quality copy or original of the letterhead.

The document containing the logo must be exported to a USB stick. See 44.9 for a description of the export to a USB stick.

On the stick you find the exported file in folder 'transfer' or 'exchange'. Open it with an image editing software (e.g. Gimp). Cut the logo out carefully and align it horizontally.

Hommel Hercules WRONG!!!!!

Hommel Hercules RIGHT!!!!!

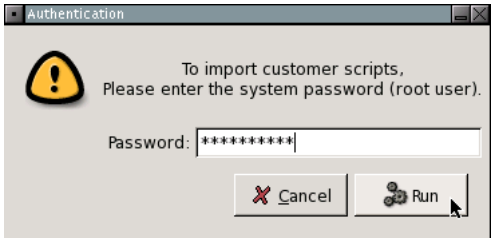
➲ Note: It is exceedingly difficult to scan a page in a way that it is absolutely straight. Nevertheless, the logo must be extremely straight because otherwise the inaccuracy produced during logo recognition regarding the rotation angle can be considerable. This in turn leads to wrong positioning of the objects. **Hence, adjust the logo so that it is in an absolutely horizontal position!**

Then, save the file in PNG format. Please do not use the JPG format because the quality of the logo would suffer.

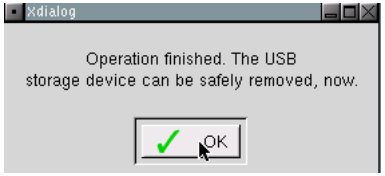
Then, take an empty (but formatted) USB stick and create a folder named 'cust'. Attach the USB stick to the ArchivistaBox. Confirm export of the 'customized' scripts.

Open the USB stick with another computer. If you do not find a directory 'formrec' inside the 'cust' folder, create such a folder (i.e. 'formrec' in directory 'cust').

Copy the logo file into the new folder 'formrec'. Copy the files back to the USB stick. ArchivistaBox recognizes the folder `cust/formrec` automatically. The following dialog box will appear on your screen:



Enter the password and click **RUN**



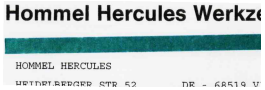
You receive confirmation that the process has been successfully completed. Click **OK**.

Now the logo in question is saved on your ArchivistaBox. In order that the logo recognition can take it as its basis you must not forget to enter the file name under menu item `Logo recognition`.

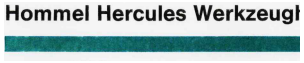
➲ Attention: In order that you do not overwrite existing logos on your ArchivistaBox you must export the 'cust' folder before you add the new logo to it and copy it back to ArchivistaBox.

22.4.3 Problematic logos from pre-printed letterheads

When you use pre-printed letterheads you must be very careful!



WRONG!!!!!



RIGHT!!!!!

The first picture shows you part of a pre-printed letterhead. All that is green is part of the pre-printed logo. What is printed in black was printed onto the sheet afterwards. As you may easily notice the printing is awry.

If you use the green bar as logo, the objects cannot be properly positioned because the information was not printed in alignment with the logo (here: bar). In such a case use that part of the letterhead that was printed later (here: 'Hommel...'); the objects may be awry but they are correctly awry in relation to the logo. The form recognition module will not have any problems to extract the objects properly.

➲ Note: It goes without saying that this procedure is only possible if there is an element in the individually printed part of the letterhead that lends itself to logo recognition. Printed script nameplates of a font size of 15pt or more as well as any frames or lines make good 'logos' in our sense of the word.

22.5 Summary

Below there is a summary of the individual steps in order that it may be easier for you to come to grips with the wealth of information in this chapter:

1. Scan form with logo
2. Export and edit file, save logo on ArchivistaBox
3. Turn on **logo recognition**
4. Define mask in **form recognition**.
5. Define object(s) in **form recognition**.
6. Set **scan definition** for form recognition. Important: 300dpi!!!
7. In general, when setting up form and logo recognition it is best to stick to absolute precision.

23 Exporting documents

23.1 Introduction

You can use the ArchivistaBox to publish existing archives at any time. This is described in detail under 38. If you only want to export specific files, certain users that can create a database from a current query in the WebClient can be specified in this form. A new (copied) database created in this way can then be published. This is useful, for example, if you provide scanning services and you want to deliver scanned documents as self-supporting archives. Alternatively, if you are a custodian, you may want to make the collected documents available to specific clients. You may also want to export an annual archive. In all of these cases, you can use the following form to make the relevant settings:

23.2 Exporting files in WebAdmin

For 'Authorised users', enter the names of all users that are authorised to call this function in the WebClient as well as the RichClient.

For 'User profile', you can choose whether the user profile of the user that is currently logged in or the user profile of a different user is to be utilised. For example, this allows you to export a database with a 'guest' user. In other words, the exported database can only be accessed using the 'guest' user account.

For 'Number of records (maximum)', you can determine the number of data records that can be exported. '0' means any number and positive numbers indicate the maximum number of files that can be exported.

You can use the 'Deactivate export' function to block the export temporarily.

23.3 Exporting files in the WebClient

After you activate the export, you can login with a previously defined user name. You can then find the menu item 'Export documents' in the WebClient under 'Choose action'. If you select this function, the system asks you to specify the required database name. When you enter a name and confirm the action, an export is carried out with the data records that are currently selected.

23.4 Exporting files in the RichClient

Under 'Database' in the RichClient, you can find the sub-item 'Export' where you can then find 'ArchivistaBox'. As with the WebClient, you are then asked for the required database name. You then also have to enter a user as well as a password so that this user can carry out the actual export.

➲: You can set up an appropriate user under 8.14.1.7.

24 Mail archiving

24.1 Introduction

You can use mail archiving to archive mail messages (Emails). You have to perform the following steps in order to do this:

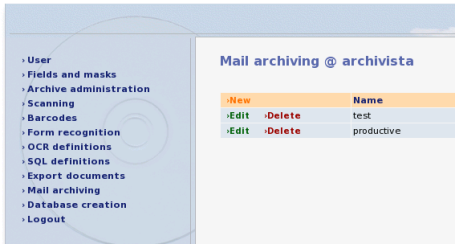
- Mail server with IMAP log (including SSL support).
- Mail archiving must be activated in the system menu (see 8.12.2 for more information).
- Mail archiving must be set up in WebAdmin for each database.
- Manual or automatic mail archiving must be completed.

This chapter only discusses the third point,that is, setting up individual mail folders that are to be archived in a database.

Before going into the individual settings options, it is necessary to point out that the following events occur during mail archiving:

- A mail message is picked up by the server.
- The mail message is analysed.
- The message itself is archived with the option of archiving all PDF files and images it contains.
- An Archivista file is opened for each mail message. The 'Subject', 'From', 'To' and 'Cc' fields as well as the date can be indexed automatically.
- The message itself is zipped and is also saved.

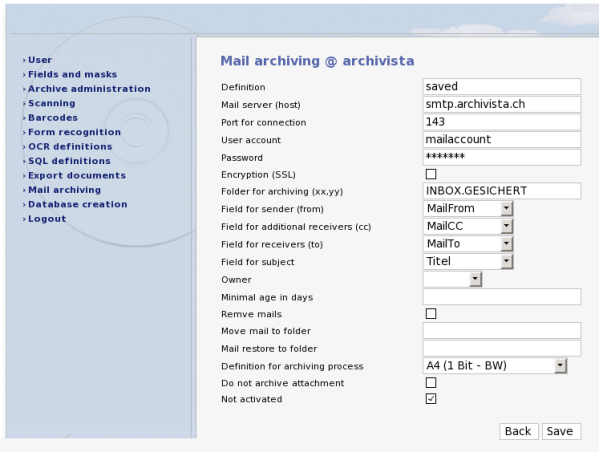
24.2 Individual definitions



If you select the menu item 'Mail archiving', the system displays all definitions for mail archiving that are active in the database you activated in WebAdmin.

If no definition exists yet, you can open one with **'New'**. You can use **'Edit'** and **'Delete'** to edit the definitions.

24.3 Editing a definition



You then have to enter a name for the **'Definition'** (this is 'saved' in our example). You should not change this name after this. Oth-

erwise it may not be possible to copy the archived mail messages back into the desired folder.

For **'Mail server (Host)', 'Port for connection', 'User account', 'Password'** and **'Encryption (SSL)'**, enter the connection data for the mail server.

In **'Folder for archiving (xx,yy)'**, you can specify multiple folders that you want to archive. Use commas to separate the entries. You cannot use any other characters to separate them.

In **'Field for sender (from)', 'Field for additional receivers (cc)', 'Field for receiver'** and **'Subject'**, you determine fields in the database that are to be used for indexing the mail messages when they are archived.

In **'Owner'**, you can specify that the archived mail messages can always be assigned to a specific user or group.

You can use **'Minimum age (in days)'** to ensure that mail messages are only archived after a certain period of time.

In future, you may be able to use **'Remove mails'** to delete mail messages after archiving. *This option has not been activated yet.*

You can use **'Move mails to folder'** to specify the folder to which a message is to be sent after archiving. This allows you to ensure that mail messages are not archived more than once.

You can use **'Restore mails to folder'** to specify the folder to which an archived message is to be returned from the WebClient when you click on the 'Mail' option in 'Download'.

You can use **'Definition for archiving process'** to use a scan definition for the post processing of mail messages. Here, you also specify whether mail messages are to be archived in black and white, grey scale or colour.

You can use **'Do not archive attachments'** to ensure that mail attachments are not added as image pages of the actual message. In each case, the whole message is archived so that it (and the attachments) can be restored at a later stage.

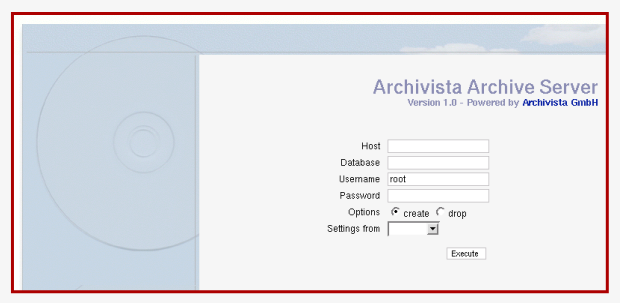
You can use **'Not activated'** to specify that a definition is temporarily inactive.

24.4 Possible problems during setup

If at first you are unable to activate mail archiving, you can find a lot of information in the log files.

You must also take into account the fact that mail archiving will only work if you activate it beforehand in the system menu under 'Mail archiving'.

25 Database creation



The screen used to create or delete archives is very similar to the login screen.

25.1 Option 'create'

If you want to create a new database, follow the steps below.

Host: enter the host name that you plan to use for Archivista 2009/I.

Under **Database** you can enter a name of your choice: it must be in one word and the initial character must not be a number.

You must enter 'root' in the **User** field if you want to create a new archive. The field **Password** must carry the root password.

The field **Settings from** gives you the possibility to 'hand down' the settings from an existing database to the new one. Select the existing database from the combo box. If you want to create a database from scratch, then you leave this field empty.

➲ Please note that you cannot delete the very first database 'archivista'.

After you click the button 'Execute' the database is created.

➲ **Important!** Please note that you can access **a new, empty** database only with user 'SYSOP'. The password is the same as for user 'Admin'. User 'Admin' must be created first before you can use it for the new database. Note too that you cannot work with the Web**Client** with user SYSOP. Hence, there is no getting round creating user 'Admin' in WebAdmin before you can enter data in the WebClient.

25.2 Option 'drop'

If you want to delete a database, then click 'drop' under **Options**.

Take care to enter the right names and to use the proper spellings when you enter data in the fields **Host** and **Database**. You cannot undo the process of deletion. Should you delete the wrong database by accident, then the only possible solution is to transfer the data from the backup medium.

Under **User** enter 'SYSOP' and the field **Password** must contain the right password.

By clicking the button 'Execute' you delete the database.

Part VI

WebConfig

26 Administration with WebConfig

26.1 Introduction

The Application WebConfig serves to undertake the webbased ArchivistaBox system settings.

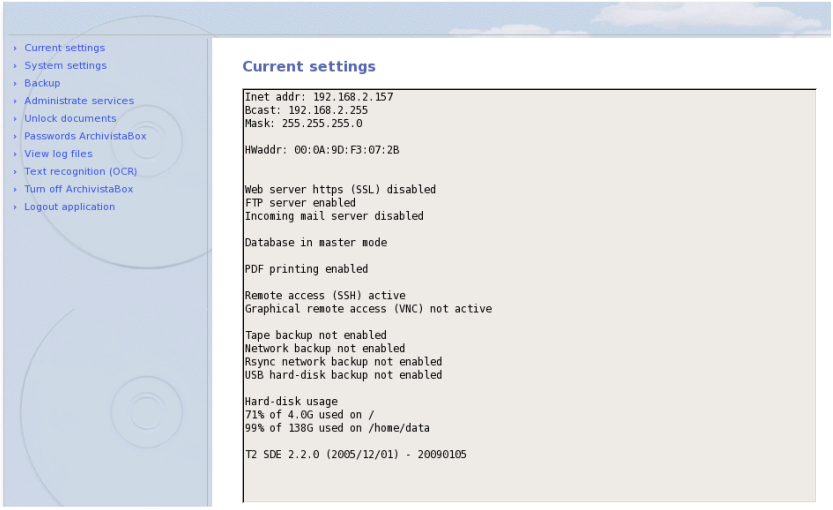
26.2 Login

In order to register you must use the password 'root'

➲ Please bear in mind that the application is possibly not available, because it is either switched off or the databank (e.g. due to Backups) is running. During this time it is not possible to do any work on Archivista WebConfig.

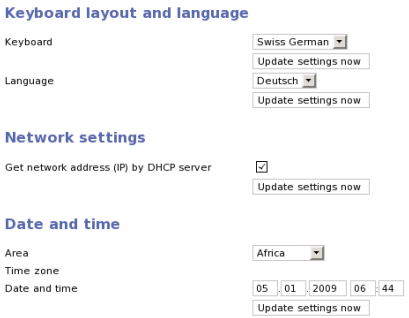
26.3 Current Settings

You will find here the current settings of ArchivistaBox.



26.3.1 System Settings

Here all settings for the keyboard, network and date, time and time zones can be updated.

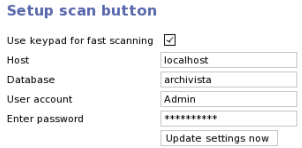


26.3.2 Mask connection WebClient

In this menu item you can define the login screen of the webclient application.

26.4 Setup scan button

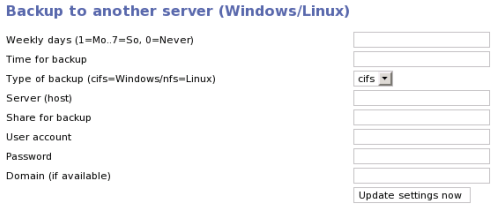
Define in this menu if the current ArchivistaBox can use the Scan button, and if so, in which database the scanned documents should be stored.

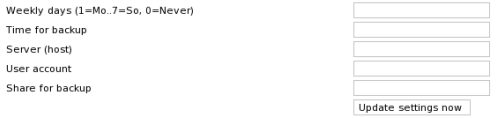


26.5 Backup

Here you can define the backup. Possibilites are:

- USB hard disk
- Network
- RSYNC server
- Tape backup

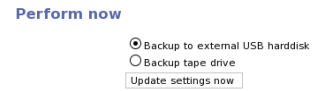




➲ Restore Data backup: For safety reasons the backup restoration is not integrated in the Applikation. Instead use the System menu under 8.8.1.3

26.5.1 Perform now

You can start a backup in this menu point point promptly.

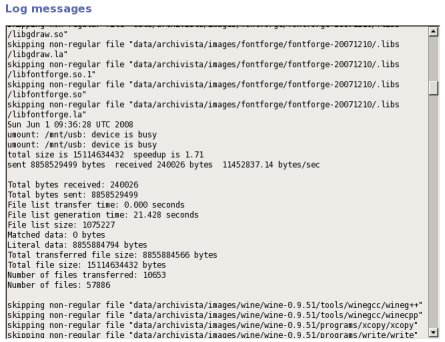


➲ **Models Dolder, Rigi, Saentis and Pilatus:** Please note that during the backup you cannot work with ArchivistaBox. Therefore ascertain that during the Backups no user is working with ArchivistaBox.

➲ **Models Titlis and Eiger:** The backup with the models Titlis and Eiger is always done on the slave box. For this reason it is guaranteed that all users can normally work on the master box.

26.5.2 Log messages

Here you will find the messages on current status of completed backups.

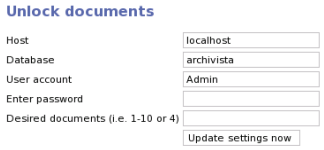

Log messages
/libgdraw.so"
skipping non-regular file "data/archivista/images/fontforge/fontforge-20071210/.libs
/libgdraw.la"
skipping non-regular file "data/archivista/images/fontforge/fontforge-20071210/.libs
/libfontforge.so.1"
skipping non-regular file "data/archivista/images/fontforge/fontforge-20071210/.libs
/libfontforge.so"
skipping non-regular file "data/archivista/images/fontforge/fontforge-20071210/.libs
/libfontforge.la"
Sun Jun 1 09:36:28 UTC 2008
umount: /mnt/usb: device is busy
umount: /mnt/usb: device is busy
total size is 15114634432 speedup is 1.71
sent 8858529499 bytes received 240026 bytes 11452837.14 bytes/sec
Total bytes received: 240026
Total bytes sent: 8858529499
File list transfer time: 0.000 seconds
File list generation time: 21.428 seconds
File list size: 1075227
Matched data: 0 bytes
Literal data: 8855884794 bytes
Total transferred file size: 8855884566 bytes
Total file size: 15114634432 bytes
Number of files transferred: 10653
Number of files: 57886
skipping non-regular file "data/archivista/images/wine/wine-0.9.51/tools/winegcc/wineg++"
skipping non-regular file "data/archivista/images/wine/wine-0.9.51/tools/winegcc/winecpp"
skipping non-regular file "data/archivista/images/wine/wine-0.9.51/programs/xcopy/xcopy"
skipping non-regular file "data/archivista/images/wine/wine-0.9.51/programs/write/write"

26.6 Services

Here you can decide whether you may print in archives, and if it is possible to deliver files by FTP upload. Furthermore, decide whether it is possible to access Archivista by remote access.

26.7 Unlock documents

Should there be a paper jam during scanning, then it could be due to the fact that a jammed document remains behind. This means that it is no longer possible to make amendments on the document, bercause the system assumes that the document is still being worked on. In this case such documents can be unlocked again using 'unblock'.

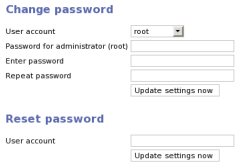


➲ Attention: an incorrect area designation can cause a queing job to be incorrectly carried out!

26.8 Passwords ArchivistaBox

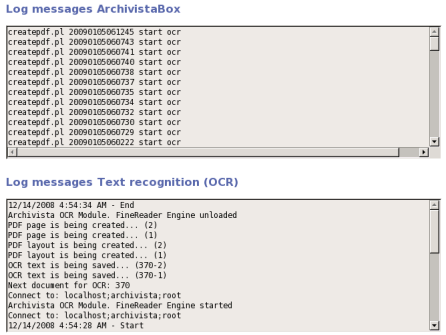
The passwords are directly administered on the ArchivistaBox. Therefore all the normal passwords can be managed directly into the WebAdmin, see also 14.

Besides, there are two user accounts, which canot be stored in the databases. Sometimes it concerns the 'root' account (owns all rights) and other times it's about the archivista user account (the user the ArchivistaBox runs under). You can define the passwords for both users here. Furthermore all user accounts for WebAdmin and WebClient can be reset. The last point you need as an Administrator, when the user has forgotten the Password.



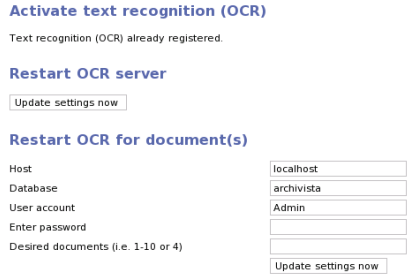
➲ Attention: Incorrectly setting these passwords as well as forgetting them makes it impossible for any future access to ArchivistaBox. Be therefore extra careful when composing these passwords.

26.9 Viewing log files


Log messages ArchivistaBox
createpdf.pl 20090105061245 start ocr
createpdf.pl 20090105060743 start ocr
createpdf.pl 20090105060741 start ocr
createpdf.pl 20090105060740 start ocr
createpdf.pl 20090105060738 start ocr
createpdf.pl 20090105060737 start ocr
createpdf.pl 20090105060735 start ocr
createpdf.pl 20090105060734 start ocr
createpdf.pl 20090105060732 start ocr
createpdf.pl 20090105060730 start ocr
createpdf.pl 20090105060729 start ocr
createpdf.pl 20090105060222 start ocr
Log messages Text recognition (OCR)
12/14/2008 4:54:34 AM - End
Archivista OCR Module. FineReader Engine unloaded
PDF page is being created... (2)
PDF page is being created... (1)
PDF layout is being created... (2)
PDF layout is being created... (1)
OCR text is being saved... (370-2)
OCR text is being saved... (370-1)
Next document for OCR: 370
Connect to: localhost;archivista;root
Archivista OCR Module. FineReader Engine started
Connect to: localhost;archivista;root
12/14/2008 4:54:28 AM - Start

26.10 Text Recognition (OCR)

In this menu teh text recognition (OCR) can be activated. Usually this should already be the case. Furthermore, you can execute the OCR-Recognition of specific files again as well as (in problems) actuate the OCR-Server afresh.



26.11 Turn off ArchivistaBox

This menu point is either used to restart ArchivistaBox, or to switch it off.

Part VII

WebERP

27 ArchivistaERP

27.1 Introduction

The ArchivistaBox originated as an DMS and archiving software. From the year 2008 Archivista will also receive an ERP Modul (Enterprise Resource Planning). We from Archivista GmbH are convinced that the linkage of an ERP with a DMS Solution on a box makes sense for various reasons.
The most important reasons are mentioned as:

- many clients ask for the possibility of an integration of their existing ERP Solution with ArchivistaBox. Dank ArchivistaERP gibt es für alle nun ein 'Anschauungsbeispiel'.
- Today everything in a box is in demand more than ever. A simple web-based ERP-Software, that can be inserted ready to use, does not exist. ArchivistaERP is therefore highly innovative.
- The number of webbased ERP-Solutions can be counted on one hand. Even in no way does OpenSource-Solutions exist like sand on the seashore. ArchivistaERP is subordinated to GPL and runs entirely webbased.
- Wer (Archivista GmbH) ourselves need a new ERP-Solution. During the evaluation we came across various products. The Solution below originates from Frontaccounting (see www.frontaccounting.com) and was adapted to our needs, so that we can work swiftly.

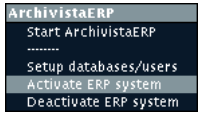
Enough about reasons, as to why ArchivistaERP exists. And so that no misunderstandings arise here, let it be said at this point that which ArchivistaERP is not:

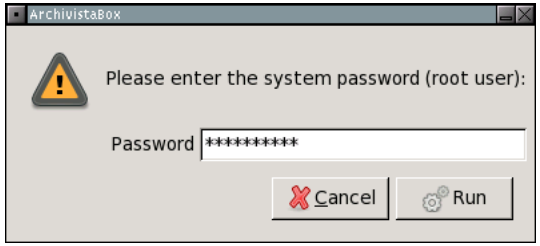
- ArchivistaERP is not a substitute for your existing ERP-solution. The crucial point with an ERP-Lösung are not in the first instance the licencing charges, but the Implementation costs. If your ERP-Solution to some extent performs the task, then please gladly keep it.
- ArchivistaERP does not (at least) at the moment offer a solution for middle and or large firms.
- ArchivistaERP is a solution among many and the ArchivistaBox supports entirely some imaginable Solutions independent of ERP-Modul with a minimum integration expenditure.
- ArchivistaERP will soon be further actively developed, i.e. in future certain adjustments must be taken into account.

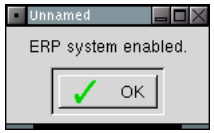
27.2 Activate ERP system

Upon delivery of the ArchivistaBox the ERP-module is not activated. For you to be able to work with the ERP-Modul, ArchivistaERP must be activated.

For that reason, the right mouse button on the Archivista desktop must be pressed. In the system menu ArchivistaERP can be activated under the Modules:







After we have activated the ERP-Solution, we can then open a databank. Select in the same menu as above the following menu:

You will arrive at the installation screen. Enter the following information:

ArchivistaERP Installation Wizard
Welcome to the ArchivistaERP Installation Wizard.
ARCHIVISTA
Please enter the database name and the root password:
Database Name: averpe
Password: **********
Company Name: Training Co.
Accounting scheme: de_CH-new.sql
Username: admin
Password: **********
Email: webmaster@archivista.ch
Re-Password: **********
Please note: ArchivistaERP is released under the GNU General Public License
By clicking install, you are accepting the license.
Install ArchivistaERP Accounting
ArchivistaERP is released under the GNU General Public License

- Database name: Name of desired Databank
- Password: root-Passwort of ArchivistaBox
- Company Name: Name of your company (amendable later)
- Database scheme: chart of accounts: (present: `de_CH-new.sql`, `en_US-new.sql`, `en_US-demo.sql`)
- EMail: Desired E-Mail-Address for admin-user
- Password/Re-Password: Passwort for the admin-user (in duplicate)

Click on 'Install ArchivistaERP'. Once all entries are completed, you will arrive at the registration screen:

You can now register yourself with 'admin' and the corresponding password. Click on 'Login'; You will arrive on the working screen of ArchivistaERP.

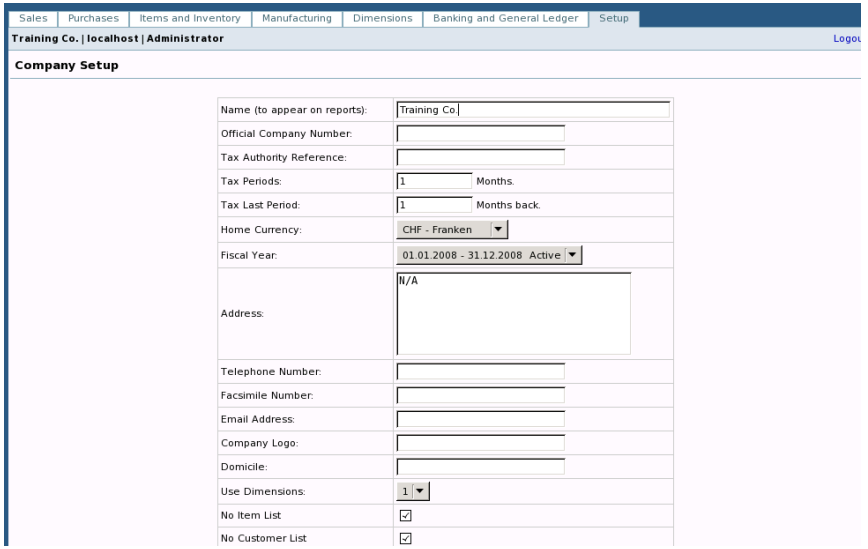
The installation is therefore complete. Sensibly you should next work through 27.3 Subsequently you should deal with a few Applications under 27.11.

27.3 First Step

In this chapter we would like to take you step by step on how to set up the first invoice. So as to enable us set up the invoice ourselves, we must next prepare the company data, at least a tax rate, a tax group, an article as well asl a further client.

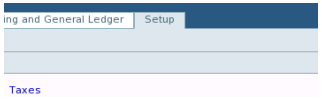
27.3.1 Company Setup

After registration change to 'Set up' as well as 'Set up Company'. You can now enter your company data:

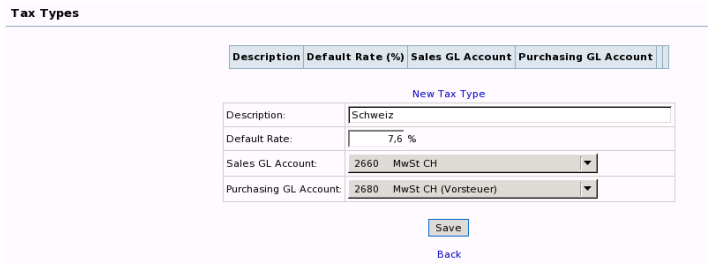


27.3.2 Taxes

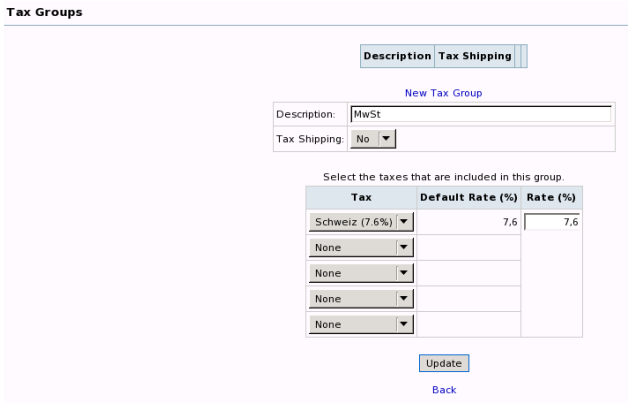
Before we can be able to compose an article, we must ascertain that there is at least one entry on: 'Tax Rate':



Make an entry there as per the diagram below:



Next to the 'Taxes' we will next need a tax group. Please click on 'Tax Groups', to open a tax group.

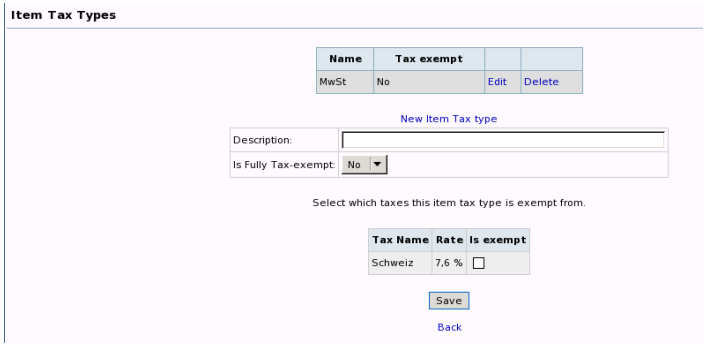


Tip: The difference between 'Taxes' and 'Tax groups' lies in that the 'Tax group' may contain several 'Taxes' and that it can be identified from the tax groups, whether or not the delivery is taxable.

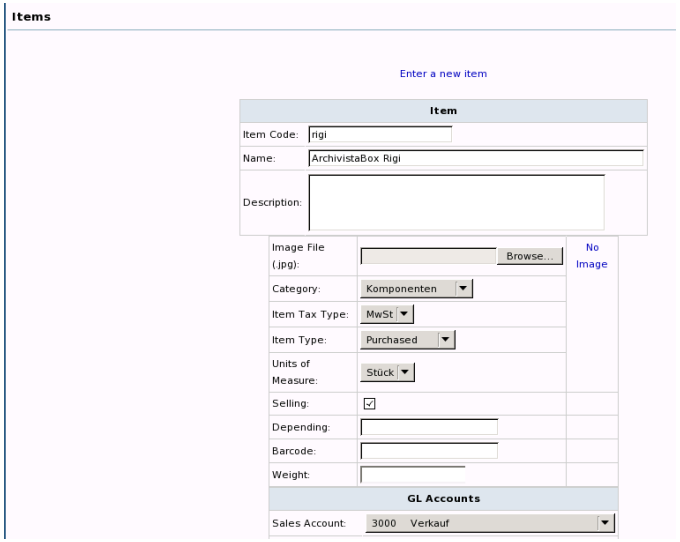
27.3.3 Manage items

Let us change to 'Items and Inventory'. Here an entry 'Item Tax Types' must be opened next:

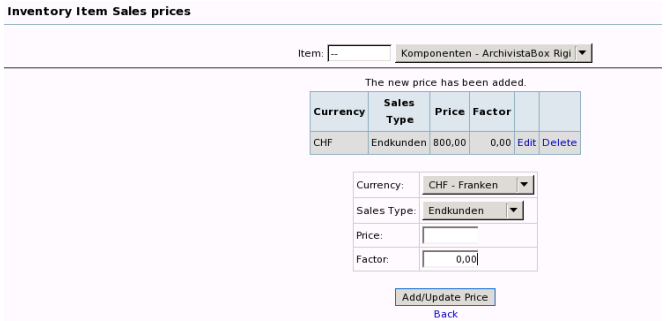
Proceed according to the graphic below:



Thereafter we can set up the first article under 'Items':

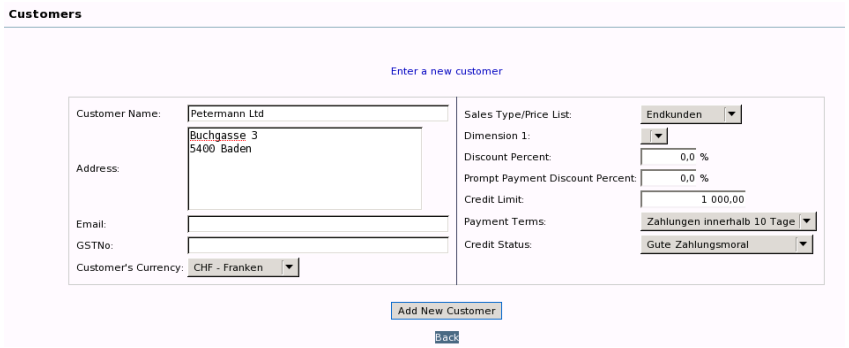


At the end click on 'Insert new Item'. Now we have compiled the first article, what is missing is that we have to attach a price on the article.



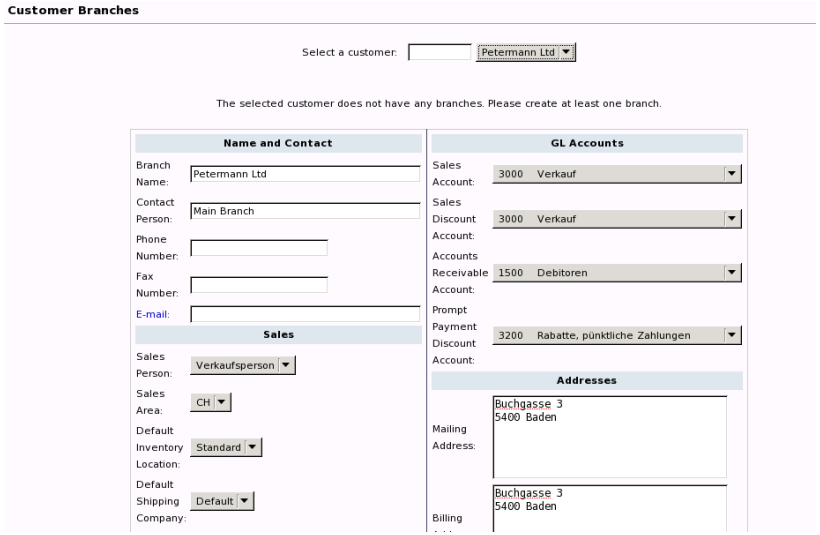
27.3.4 Add a customer

So as to be able to commence with the sales, we must open a customer account. Go to 'Sales' and then to 'Add and Manage Customers'. Enter similar data in the field:



Important: The address should always be entered in a memo field in ERP. Thereby you decide at the same time the formating of the address.

In order to open a customer account, click on 'Add new customer'. Additionally, for every client, we must open a branch. You will find this under 'Add branches for this customer':

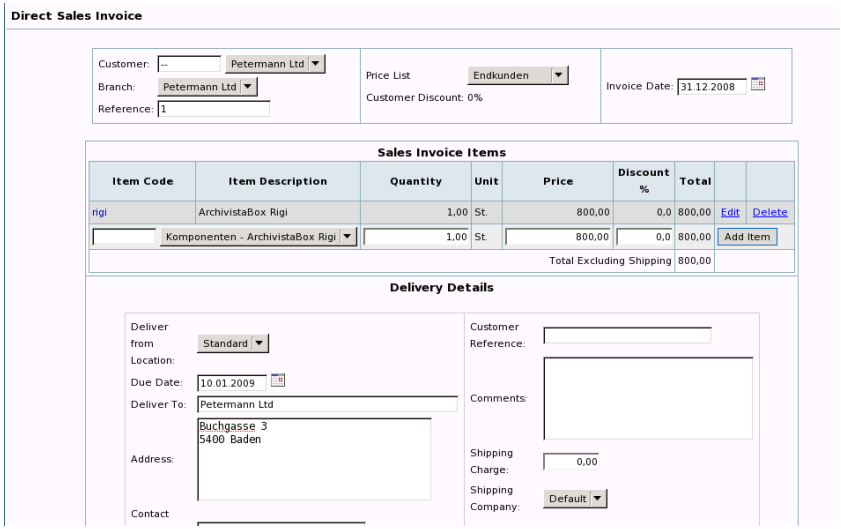


Please note that every branch will need a new address where the the invoice and delivery can be entered. In our case, we take over the address details and of the company.

27.4 Create invoice

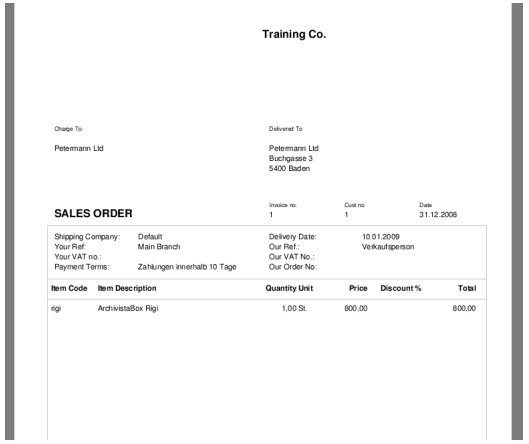
In order to compile an invoice, choose 'Direct invoice' at 'Sales':

In order to add an article, either enter the article code or (here rigi) use the drop down option to choose from the list. So that the article can be charged on the system, finally click on 'Add item'.



The invoice will be set up by clicking 'Place invoice'. Due to the fact that some information must be updated, it can happen that the button 'Place invoice' must be accessed twice.

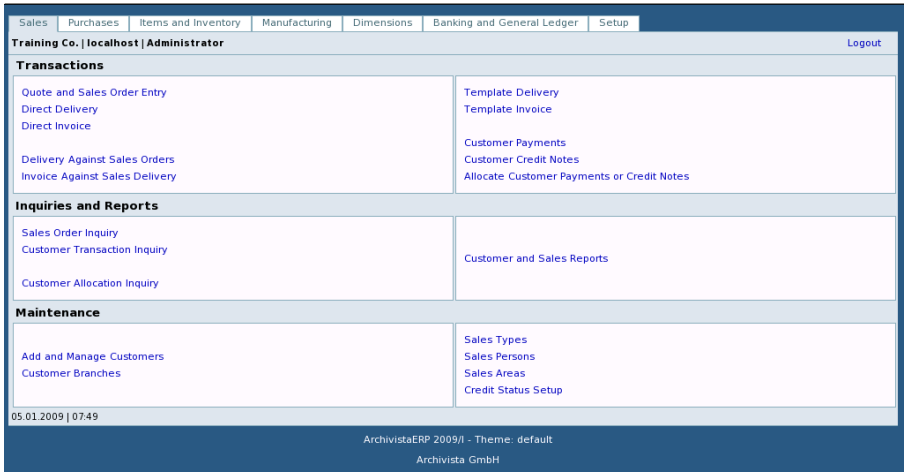
The invoice is created and already booked too. It can be printed by pressing 'Print sales invoice' as a PDF. The result should look something like this:



➲ Please note that booked invoices can no longer be deleted. You should realise this right from the start. We there really recommend that you initially experiment with a sample mandate, before the irrevocable version is 'finally entered'. A new company can be entered as described under 27.2.

27.5 Sales

In the sales section all offers, orders and invoices can be handled. Likewise, the administering of clinet addresses are found here.



27.6 Purchases

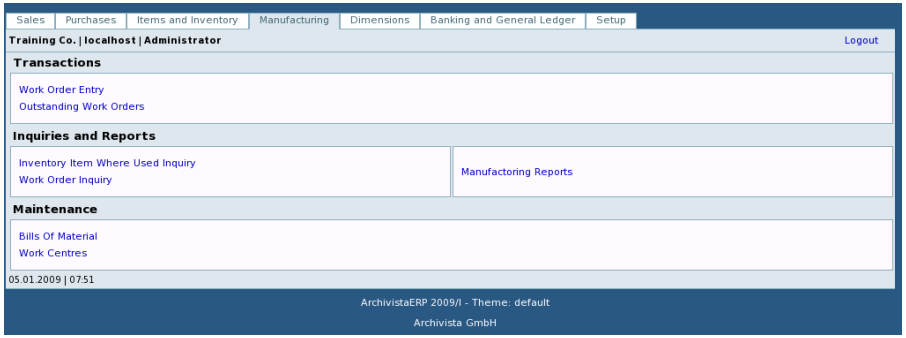
Under Purchase the purchase orders are dealt with. To this end belong also the delivery administration.

27.7 Items and Inventory

In this menu we administer the articles just like in the inventory section.

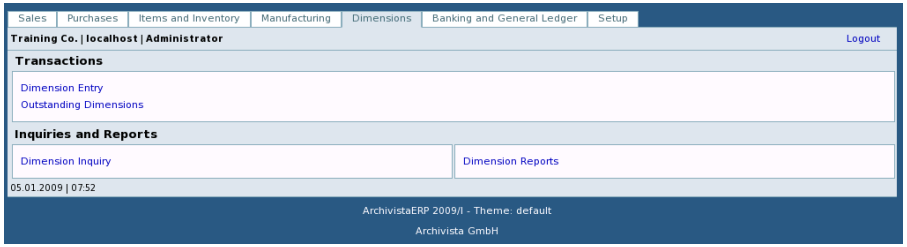
27.8 Manufacturing

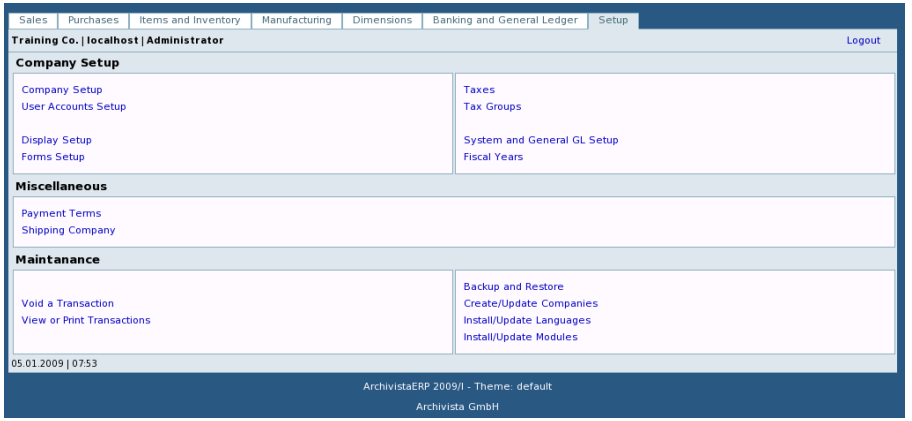
Manufacturing planning can be carried out here.



27.9 Dimensions

Under Dimensions we find cost centres and/or profit Centers.





27.10 Banking and General Ledger

Here you will find the bookkeeping as well as all options for payments and bank accounts.

27.11 Setup

In this main menu diverse options can be preset.

Part VIII

RichClient

28 First steps

28.1 Before you begin...

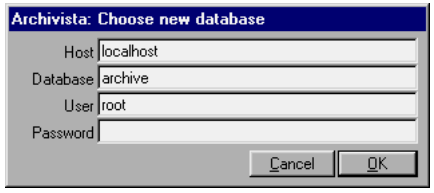
In the following subchapters we show you in step-by-step fashion how you start up the program, create new documents and add pages. We proceed on the assumption that Archivista 2009/I is installed and the first database is in place. If this is not the case, please work through chapter 'Introduction' and subchapter 'Installation'.

➲ Please take note of the online tutorials, which show you in movie-like fashion how Archivista works. You will find the online tutorials on the CD or on our web pages www.archivista.ch.

28.2 Starting up the program

Start the program by clicking the 'Start' button and pointing to 'Programs'. You will find the entry **'Archivista 2009/I (e)'** and within the Archivista product range the program 'Archivista 2009/I'. Click on this and the program starts up. Depending on the version you have or, respectively, depending on whether any users have been created already you log into the program either with or without login dialog box.

➲ If you get a screen asking for your password, please enter host (server), database, user name and password. You get this information from the systems administrator.

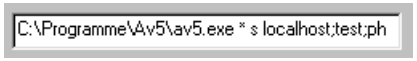


28.2.1 Desktop link to the database

If you create a desktop link to the program, you can go a step further and preset host name, database name and user. How to go about it:

- Create desktop link: click on **'Start'** and move the mouse to 'Programs' and 'Archivista'. Create the icon by means of a right click and 'Send to'.

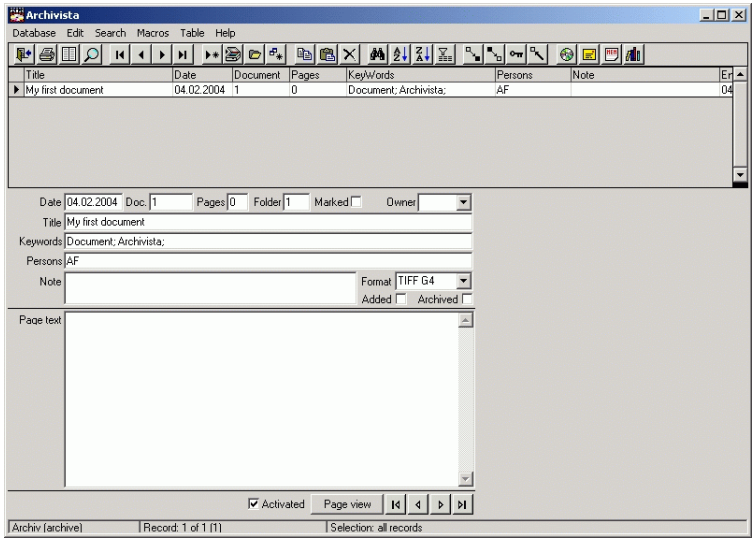
- Rightclick the new icon and complement the target under 'Properties'. The example below gives 'localhost' for host, 'test' for the database name and 'ph' for the user.



28.3 Creating a document

Tutorial on Creating a Document (movie 2min) *5.1*

After starting up you get to the table 'Archive'. Below an illustration of this table. Click the icon in the toolbar and the document is being created – at the same time the system automatically fills in the current date in the field 'Date'.

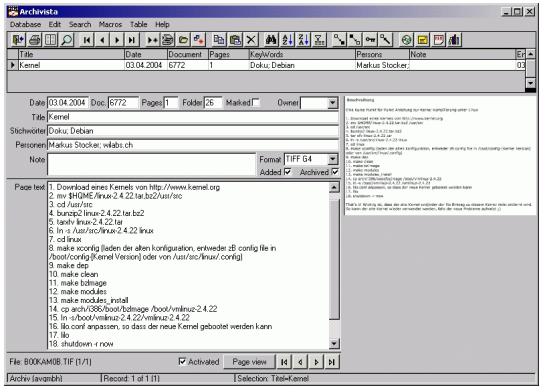

Archivista
Database Edit Search Macros Table Help
Title Date Document Pages KeyWords Persons Note
My first document 04.02.2004 1 0 Document; Archivista; AF
Date 04.02.2004 Doc. 1 Pages 0 Folder 1 Marked Owner
Title My first document
Keywords Document; Archivista;
Persons AF
Note
Format TIFF G4
Added Archived
Page text
Activated Page view
Archiv (archive) Record: 1 of 1 (1) Selection: all records

Please make your own **entries** now. Taking the screen above as example enter the fields 'Title', 'Key wrds' (Keywords) and 'Persons'. To do that click the relevant text box with your left mouse key. Your entries will be saved automatically. You must not change the fields 'Document', 'Pages' or 'Folder'. The document is now ready for you to add pages.

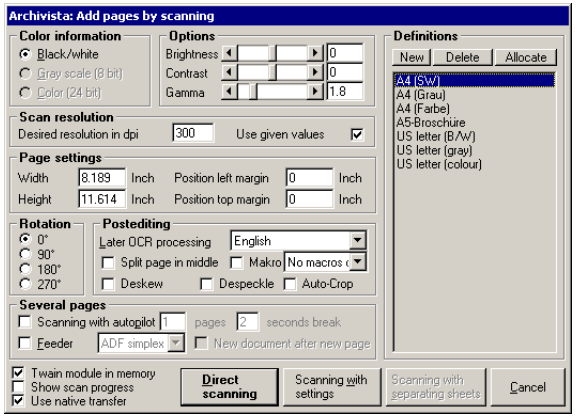
28.4 Adding pages

Tutorial on Adding pages (movie 2min) *5.1*

To add pages you must change to another view. Press the button Page view and you get to the **'Page view' mode**. You achieve the same result by pressing the function key **'F9'.**

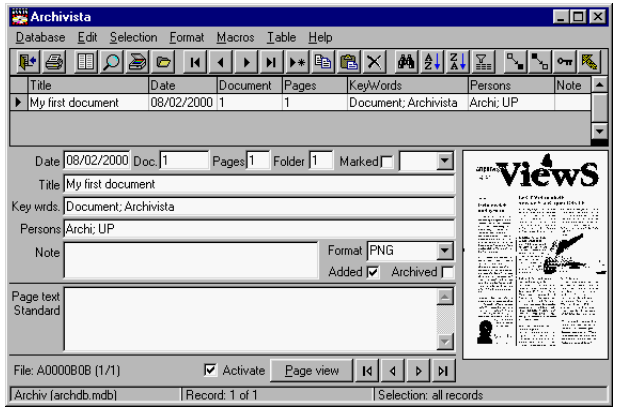

Archivista
Database Edit Search Macros Table Help
Title Date Document Pages KeyWords Persons Note
Kernel 03.04.2004 6772 1 Doku; Debian Markus Stocker;
Date 03.04.2004 Doc. 6772 Pages 1 Folder 26 Marked Owner
Title Kernel
Stichwörter Doku; Debian
Personen Markus Stocker; wilabs.ch
Note
Format TIFF G4
Added Archived
Page text
1. Download eines Kernels von http://www.kernel.org
2. mv $HOME/linux-2.4.22.tar.bz2 /usr/src
3. cd /usr/src
4. bunzip2 linux-2.4.22.tar.bz2
5. tarxfv linux-2.4.22.tar
6. ln -s /usr/src/linux-2.4.22 linux
7. cd linux
8. make xconfig (laden der alten konfiguration, entweder zB config file in /boot/config-[Kernel Version] oder von /usr/src/linux/.config)
9. make dep
10. make clean
11. make bzImage
12. make modules
13. make modules_install
14. cp arch/i386/boot/bzImage /boot/vmlinuz-2.4.22
15. ln -s/boot/vmlinuz-2.4.22/vmlinuz-2.4.22
16. lilo.conf anpassen, so dass der neue Kernel gebootet werden kann
17. lilo
18. shutdown -r now
File: BOOKAM08.TIF (1/1) Activated Page view
Archiv (avambh) Record: 1 of 1 (1) Selection: Titel=Kernel

The page is empty since there has not been anything scanned yet. Now, click to scan a page. If there is no scanner installed, go to menu 'Page' and menu item 'Import' and then to 'From file..'. After entering the file name the page is imported. If you want to add pages with the help of a scanner and in the case of such a device being installed correctly, you get the following form:



To add a page which does not contain any information on color or gray scales, click on 'US letter (B/W)' under 'Definitions'. After that choose **'Direct scanning'.** The scanning process is started and the page is photographed and put into a file.

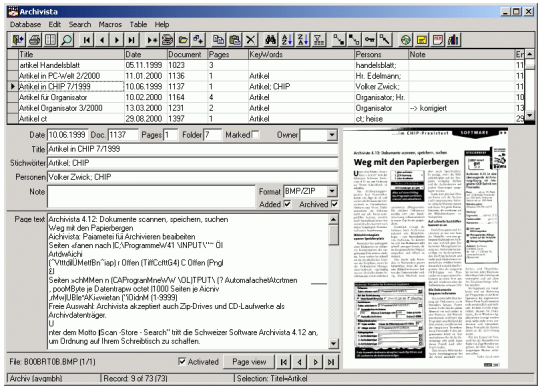
After the page is added it is displayed in its full size. To see it enlarged use the magnifying glass. It goes without saying that a variety of other functions is at your disposal. You do best if you try out the items of menu 'Edit' right now. To get back to the table mode of table 'Archive' you must leave 'Page view'. Press **function key 'F9'** to do that.



The added page appears now also in the standard view of table 'Archive'. The field 'Pages' was automatically updated. This is the end of 'first steps'. We recommend you to create a few more documents and pages before you go to other chapters in this manual.

28.5 The main view of table 'Archive'

After starting you will always get to the main view first, i.e. the table mode of table 'Archive'. Let us take a look:



The working space of Archivista can be roughly divided in four areas:

- **Area functions** with menus and toolbar (top part of screen)
- **List (of records)**
- **Detail view**
- **Status bar**

The **functions** (i.e. the menus and the toolbar) are needed when you want to process a certain command (when creating a new record, for example). Please note that each existing function must be part of a menu but does not necessarily exist as an icon in the toolbar. The menu items are described in more detail in the following chapters.

You need the **list** above all to browse through the records click the list, then use the arrow keys to activate other records and fields.

You can also make entries in the list. Press the Enter key to reach the editing mode. Please note, too, that the list contains the data of table 'Archive', but that table 'Archive' encompasses more than the list.

➲ Only one record can be active at a time. You recognize the active record by a black triangle on the left hand side of the list.

The lower half of the screen is reserved for the **detail view.** On the left there are the details of the active document and the page text if the OCR process has been run. In the lower left hand corner the **file name of the page** currently displayed is shown (without extension). On the right you see **a page of the active document** (only if a page has been added already). With the button **'Page view'** you can change to an enlarged view of the same page where additional functions await you. To the right of the button 'Page view' there are some **buttons to navigate** among the pages of the active document.

The **status bar** informs you of certain basics of the currently displayed table. On the left the name of the current table appears, in the middle the name of the active record as well as the total number of records of the current selection. On the right the active selection is shown.

➲ By selection we mean a certain part of the entirety of records chosen on the basis of a feature that is common to these records (see menu 'Selection').

Please note too that although this is the table mode of table 'Archive' and the main view, table 'Archive' encompasses more such as 'Page view' and the report generator.

28.6 Further tables of Archivista

Beside **table 'Archive'** there are three main tables, namely tables **'Addresses', 'Notes' and 'Literature'.** Furthermore, there are a few auxiliary tables. See chapter 34.

29 Menus, toolbar and shortcuts

29.1 Menus and functions

All functions (=what you use when you want to make something happen) at your disposal in Archivista 2009/I are to be found in the **Menus** (second bar in the program window).

A detailed description of the menus follows in the next chapters. Here only a short overview is given.

In menu **'Database'** you will find those commands that are valid for the whole database. With the commands in menu **'Edit'** you make changes on the level of the individual documents or you activate a specific document. In menu **'Selection'** you can filter your records in order to work only with a specified range of them (all documents created before 1 Jan 1990, for example). Menu **'Format'** features formatting variants for the memo fields. Menu **'Macros'** serves to run previously defined macros. Finally, in menu **'Table'** you can select the table you want to work with (i.e. you can change to 'Addresses', for example) and under **'Help'** you receive additional information on the program.

29.2 Toolbar

You will find the most important commands in the toolbar. When you move your mouse pointer to a specific icon and hover over it for a moment you will get a tiny window with a short description of the function.

29.3 Shortcuts

Almost all functions in the menus can be run by a shortcut. You can always find the key combination for the shortcut on the right hand side of the menu command. Working with shortcuts is very efficient, this applies to all Windows programs not only to Archivista. The following descriptions of the menu commands give the shortcuts in brackets.

29.4 Tab key

The **'Tab key'** (i.e. the key on the upper left side with the two opposed arrows) serves for easy navigation within a record. With this key you can comfortably move from one field to the next. To move one field backwards you press the combination **'Shift+Tab'**.

29.5 Resizing the detail view window

You can enlarge both the page (graphic) as well as the field with the recognized ASCII text on the left (field 'page text', only containing text if you used the option 'OCR' when archiving). Click the **black rectangle** on the right side under the table, **keep the left mouse key pressed** and move your mouse upwards. You will notice that a black **bar** appears on the screen. This bar marks the new relation of table to detail view. As soon as you **let go of the left mouse key** the page and the field with the ASCII text are bigger in size. In the same manner – by pulling the black rectangle downwards – you can effect a smaller graphic and a smaller field 'page text'.

29.6 Using the right mouse button

Within Archivista 2009/I you can **click the right mouse key** and you will get a shortcut menu with appropriate commands.

29.7 Special fields

In each of the four main tables there are three special fields. Since we already know the general meaning of fields we shall immediately go on to the specific use of the special fields.

29.7.1 Field 'marked'

The field 'marked' (see illustration above left or detail view above right in table 'Archive') is a so-called **'Yes/No' field,** i.e. either the record is marked or it is not marked. You can use this field primarily to make a fast selection of a few records: click this field (and make a small x appear) in all desired records. Afterwards go to menu 'Selection' and menu item 'marked'. Now the list shows only those records which you marked before, i.e. those containing a small x in the field 'marked'. Now you can work with these records. You can print them, for example, or you can search them (with 'Archive search mode'). Do not forget to delete the check mark of the records later when the specific selection you made does not make any sense any longer.

29.7.2 Field 'owner'

➲ You need this field only if the same database serves several users. You can omit reading this paragraph if this does not apply to your case.

The field 'owner' (illustration above, on the right) allows you to allocate each record to exactly one user, i.e. a certain record belongs to that specific user. Click the pull-down triangle on the right side of the field and a list with all active users appears. Now, choose the desired user. The user 'ALL' means that the record can be accessed by all users (=does not belong to one owner in particular).

If a document is allocated to one user and the others have no right to access this document, then, obviously, the other users cannot view nor work on that particular document. By using the field 'owner' you can effect that certain documents are protected because it may well be that the director does not want her employees to have access to payroll information.

29.7.3 Field 'counter'

This field is **not shown in detail view.** However, you can see it if you left-click the list and use **'the right arrow key'** to move to the last field. The last field is the field 'counter'. It is called like that because whenever you create a new record this field will automatically show the current number of records, i.e. an unequivocal value which is kept throughout the lifetime of the record and cannot be edited.

29.8 Additional fields in table 'Archive'

For table 'Archive' you can define up to ten additional fields. When you start Archivista they will be automatically added to the detail view of the table mode.

To make such modifications go to the submenu **'Define fields'** (see menus 'Database' and 'Parameters'). Further information on adding fields is available in chapter 36.3.

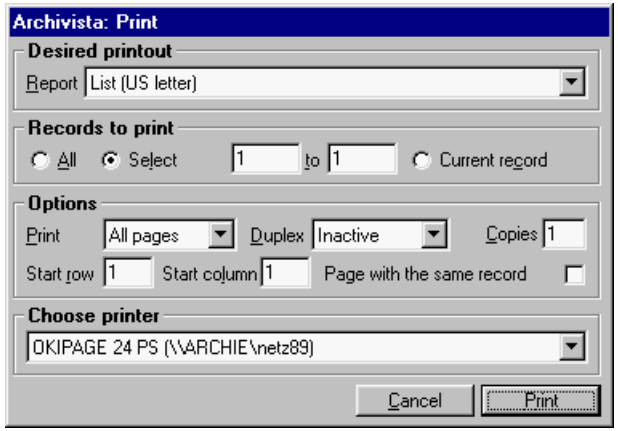
30 Menu Database

In menu 'Database' you find those commands which apply to the whole archive. In the following you find the descriptions of the individual menu items one by one.

30.1 Printing (Ctrl+P)

Select menu item 'Print...' to convey your data to paper. Printing from the main view means printing with the help of print reports (lists, labels, etc.).

➲ If you want to print individual pages, please go to 'Page view' and run the printing command from there.



Choose the kind of printout you wish under **'Desired printout'** by choosing a **'Report'.** By (print) report we understand a definition to print out data in a certain layout. Archivista contains a certain number of pre-defined print reports. In addition to these you can create your own reports by using the module 'Print reports'.

Further down you specify which records you want to print **('Records to print').** You can choose all records or a selection of them.

Under **'Options / Print'** you specify whether you want to print all pages, all odd or all even pages. What is more, you can determine the number of **'Copies'** per page. **'Duplex'** gives you the option of printing single-sided or double-sided documents (only available for printers supporting duplex printing).

The fields **'Start row'** and **'Start column'** are only needed in the case of printing labels by using sheets of labeling paper which are already partly printed on. Archivista always prints in rows, i.e. a row is finished first (covering all columns) and then the next row is begun. To begin a printout in the third row second column, please enter the value '3' after 'Start row' and '2' after 'Start column'.

If you tick **'Page with the same record'**, then a whole page is printed by repeating the same record again and again (i.e. with labels having 8 rows and 3 columns you will get the same record 24 times). This option is useful to comfortably replenish your stock of address labels belonging to a certain record.

Finally you specify the desired printer with the field **'Choose printer'.** Here Windows standard definitions are used.

When you are finished with the details click **'Print'**. A status message follows telling you which record is being printed. If you wish to interrupt the current printing process, click **'Cancel'.**

30.2 Working with databases

30.2.1 Logging on with password

This function allows you to log on afresh without having to quit the program. This is a considerable advantage when different people work on the same computer and have to make entries in an Archivista database.

➲ A password is only required if users have been defined previously (see chapter 35.2).

30.2.2 Choosing another database

With this menu item (and the relevant submenus) you can get straight into one of the four databases most recently looked into.

➲ As long as you work with only one (the first) database, you will not need these menu items. On start-up of the program you will always get into the database with which you worked the previous time.

30.3 Importing and exporting documents

30.3.1 Exporting data

It goes without saying that Archivista 2009/I allows you to export your data. At this point in time you can export your records in the formats ASCII and HTML (the latter only from table 'Archive'). By using the ASCII format virtually all requirements regarding export are met.

The export of records is very simple indeed. Select the documents that you want to export, which goes to say that the export function always pertains to the current selection. Please note, too, that an export can take place only to an empty directory.

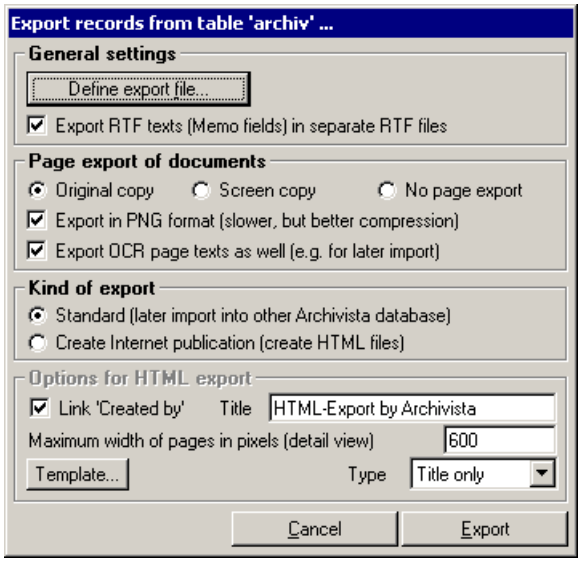
Run the menu item 'Export data...'. In the dialog box that appears you must first define the export file. If you want to export the **RTF or the memo fields** as separate RTF files per record, you must choose the corresponding option.

The export itself may take quite some time and depends on how much data you want to transfer. Cancelling the export is possible at all stages. After the export is completed you will find all the relevant files in the directory that you specified earlier.

Records that have been exported (only export in ASCII format) can be imported into any other Archivista database. Furthermore they can be imported into other programs. In both cases take heed of the fact that the first line of the export file contains the field names of the table while the lines that follow contain a record each. The individual entries, i.e. the fields or field definitions are separated by tab stops.

30.3.1.1 Exporting pages of table 'Archive'

When exporting from table **'Archive'** you must further specify if you want to export the **pages** too and if you want to export them as 'original copy', 'screen copy' or as a HTML file.



Further you must decide whether you want to export the **OCR text pages in ASCII** (this option does only make sense in those cases where the documents were previously treated with menu item 'OCR').

➲ When exporting pages please make sure that the target drive offers enough space for all of them. Otherwise the export will not be completed due to a lack of disk space.

30.3.1.2 HTML export in table 'Archive'

HTML export creates a ready-to-read HTML publication. On the main page you will find all the information of the individual documents and on the subsidiary pages the individual pages will be shown in screen copy resolution (compressed as JPEG file). Navigation occurs in the first line of the subsidiary pages whereby [Home]=Go to main page, [<<<]=previous document, [<<]=first page of document, [<]=previous page, [>]=next page, [>>]=last page of document and [>>>]=next document.

➲ Please note that the HTML export function always covers the current selection, i.e. you must define which documents you want to export and turn into a HTML publication BEFORE the actual export.

There are a number of settings which enable you to customize the look of your HTML publication.

Created link: You can switch database names and 'Created with Archivista' (for Archivista Business) on or off, as preferred.

The **title** can be specified individually for each HTML publication.

Maximum width: Use this parameter to specify the width of the bitmaps within the single-page view. This means you can easily fit your HTML publications into existing home pages (e.g. frames).

Template: You have the option of using a template file. Please note the instructions in the 'avexpo.htm' file.

Type: Select the HTML publication type. 'Detailed' will create a publication with all fields, 'Title only' means that only the title line will be exported, and with 'Photo album' only reduced-scale thumbnail pictures will be created.

30.3.2 Importing data

The menu item 'Import data...' is the companion piece to 'Export data...'. You can very easily import data that has been exported earlier from another Archivista database. Just import the file previously defined as export file (Archivista parameter file). When running 'Import data...' you will be asked for such a file. Click 'OK' to import the corresponding data. When importing into **table 'Archive'** you must further decide whether you also want to import the pages that belong to the documents.

30.4 Hyperlinks (links between records)

You can link records of the tables 'Addresses', 'Notes', 'Archive' and 'Literature', i.e. you can comfortably jump from one record to the next although they may not be close in the list. The following four functions also exist as icons in the toolbar (on the right).

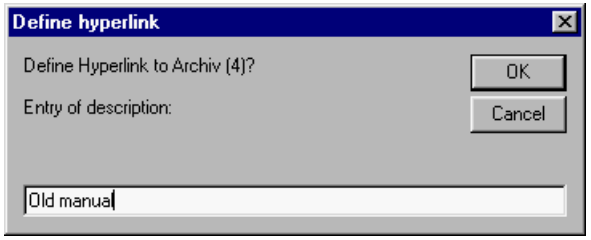
30.4.1 Defining a target (Ctrl+I)

To create a link first go to the record you would like to jump to (from another record). By first moving the cursor to this record and afterwards running the function 'define target' (either via the menu or the icon) this record is defined as the target record.

➲ When creating links in table 'Archive' the current page is included. Thus the target is a specific page in a document. This allows you to comfortably jump to a specific page and to jump from one page to the other within the same document.

30.4.2 Defining hyperlinks (Ctrl+E)

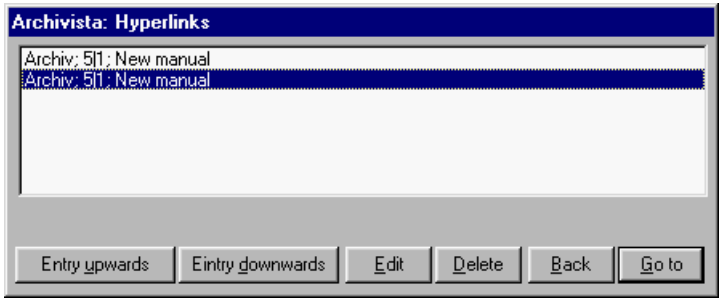
After defining the target you define the link by determining the source. Go to the record from which you want to jump. Now click 'define hyperlink' and you will get the following screen:



Enter the name of the hyperlink (if you like you can accept the suggestion made by Archivista) and click **'OK'.** The hyperlink is defined. You can tell this by the **key** in the toolbar that is now shown **in red.**

30.4.3 Call up hyperlinks (Ctrl+F8)

The current hyperlinks of the active document/address item etc. can be looked at by clicking the red key or by choosing 'Current hyperlinks' in menu item 'hyperlinks' (menu 'Database'). Below an illustration of the list that will appear. Click the desired entry and you will be carried to the target record.



30.4.4 Last active record (F8)

By running this function you get to the previously active record. It may be the record from which you jumped to the present one by means of a link or it may be the last active one.

30.5 Parameters

With this menu item you can tailor Archivista 2009/I to your individual needs. Please go to the separate chapter 35 further towards the back of this manual.

30.6 Page view (F9)

The menu item **'Page view'** enables you to change to a view where you can magnify pages and edit them or where you add them in the first place. Please note separate chapter 33.

30.7 OCR settings current page

With this submenu item you determine the details of how the text recognition software treats the currently active page. You must differentiate between 'Standard' and 'Barcode' recognition as well as between 'Exclude' and 'OCR done'.

The **OCR defnitions** serve to 'tell' the text recognition software what symbols on the page it should go for and in what way the found symbols should be interpreted. Up to 20 OCR definitions can be specified (see chapter 33). The entries you make there will appear in 'OCR settings current page'. Each individual page can be allocated a definition this way. The first OCR definition may be called up by the shortcut 'Ctrl+X'.

If you choose 'Exclude (Ctrl+Y)' the current page will not be treated by the subsequent text recognition. If you activate 'OCR done (Ctrl+Z)' the current page will not be treated by the text recognition either but you can enter your own text in the field 'page text'.

➲ The menu 'OCR settings current page' is to be found in table mode of table 'Archive' in menu 'Database', submenu 'Archiving' as well as in 'Page view' in menu 'Page'. If you click your right

mouse button when you are in table mode and pointing to the words 'Page text' in the detail view 'OCR settings current page' will also appear as context menu.

30.7.1 Extending OCR settings to entire document or selection

The OCR settings of the currently activated page may be made over to the entire document or even the entire selection of documents. By one stroke you can therefore specify, for example, if you want to subject one or several documents already treated by text recognition to a second text recognition. When at a later stage the internal OCR is run, these documents are treated afresh (with different parameters if you wish).

30.8 Quitting (Alt+F4)

The menu item **'Quit'** quits the program. A confirmation mask appears which must be acknowledged with the **'Return'** key. You can leave the program also by using the shortcut **'Alt+F4'**

31 Menu 'Edit'

Tutorial on Menu 'Edit' (movie 2min) *5.2.3*

In menu **'Edit'** you find all commands to move efficiently within a table, to edit individual records, to delete and copy them or create new ones.

➲ The commands in this menu always apply to the selected records, i.e. those records visible in the list. If you want to be sure to include all records, press 'F6' before running the command (or go to menu 'Selection' / 'All records').

31.1 Working with the clipboard

31.1.1 Marked parts on clipboard (Ctrl+Ins)

Marked text bits can be copied onto the clipboard. The text is put on the Windows Clipboard and can be recalled later either by another program or by Archivista.

You **mark text bits** by using the mouse (mark the beginning of the bit you want to copy by a left-click, keep the mouse key down, move to the end of what you want to copy and let go of the mouse key). You can also mark text pieces by the key combination **'Shift + Arrow keys'!**

31.1.2 Cutting text and moving it to clipboard (Shift+Delete)

Contrary to the command 'Marked parts onto clipboard', which does not change the marked bits, this command removes the marked pieces. At the same time they are being put on the clipboard.

31.1.3 Pasting text from clipboard (Shift+Ins)

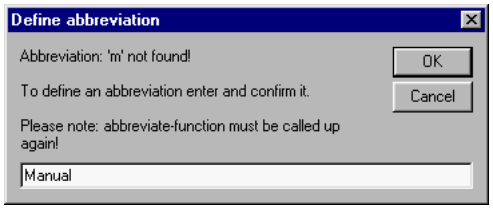
This menu item captures the text previously copied onto the clipboard and inserts it where the cursor is situated at the moment, i.e. into the active field of the active table.

31.2 Working with abbreviations

Tutorial on Working with abbreviations (movie 2min) *5.2.2*

31.2.1 Abbreviations (F2)

In Archivista you can define abbreviations in order that you may not have to type in the same words again and again. Calling up the abbreviation is simple: when you are in the field in which you want to enter an often used word, type the abbreviation and press **'F2'.** Archivista will replace the abbreviation by the previously defined term. This works only if an abbreviation was defined in the first place. If there was not, the following mask appears:



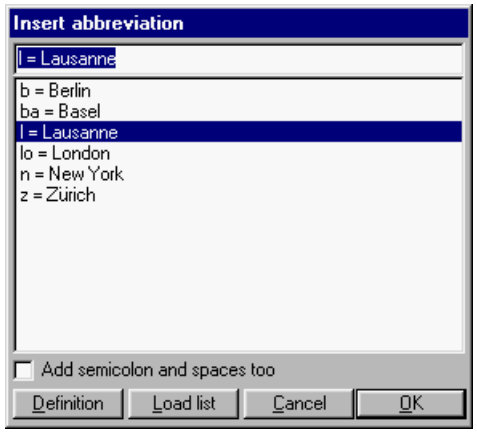
You have now the possibility to define the abbreviation. Enter the full term that you want to link with the abbreviation and press **'OK'.** The abbreviation is now defined and from now on you can call it up any time by pressing **'F2'**.

➲ The shortcut 'F2' enables you to call up any previously defined abbreviation no matter where you are. In addition to that you

may edit the abbreviations by going to the auxiliary table 'Abbreviations'.

31.2.2 Get abbreviations window (Ctrl+F2)

If you are no longer familiar with your abbreviations you can get a list of the already defined abbreviations by running this function. You will get the following window:



With a mouseclick on the desired expression and a confirming 'OK' the previously defined text is entered in the current field. When you have a large number of abbreviations you can enter the first letter of the desired word and you will be carried automatically to the words beginning with this letter.

Note: The option 'Add semicolon and spaces too' helps you to keep keywords separate if you enter several of them. The option 'Load list' serves to bring the list up-to-date after you have defined abbreviations by means of 'F2'.

31.3 Navigating among the documents

31.3.1 Activate first record (Ctrl+F3)

With this command the cursor (the little black triangle on the left) moves to the first record within the current table and the current selection. The record will be visible in the detail view and can be edited.

31.3.2 Go to previous record (F3)

With this command the record immediately preceding the currently active one is activated. In most cases you will observe how the little black triangle on the left of the active record moves to the record just above it (unless it is the first record in the list).

31.3.3 Go to next record (F4)

In this case the cursor moves to the record just below the currently active one.

31.3.4 Activate last record (Ctrl+F4)

With this command the cursor moves to the last record within the current table or the current selection, respectively.

31.4 Create, copy, paste and delete documents

31.4.1 Create new document (Ctrl+N)

Tutorial on Create new Document (movie 2min) *5.2.4*

By running this command you add a new record to the table (e.g. a new document to the table 'Archive'). You can run this command only if you possess the right sort of user privileges.

➲ Archivista supports four different archiving formats: 'Encrypted (Bmp/Zip)', '1 Bit Black/White (Tiff G4)', 'Loseless (Png)' and 'Compressed (JPeg)'. From version 2009/I onwards the archiving format may be specified for each individual document by choosing the corresponding option in field 'Format' before adding the first page. The default value when creating a new document is 'PNG'. This default may be changed in 'Parameters'/'Archive'/'Archiving format'.

31.4.2 Copying a document (Ctrl+C)

The whole record is copied onto the internal clipboard and can be retrieved by the command 'Paste record' (see below).

Note: In contrast to the menu item 'Marked parts on clipboard', which refers only to a part of a field, the menu item 'Copy record' comprises the whole record.

31.4.3 Paste document (Ctrl+V)

A record which was previously put onto the internal clipboard with the command 'Copy record' is recalled with 'Paste record'. Out of the existing data a new record is created. For this new document the current date is taken, the document number is increased by one and the field 'Pages' contains a '0' since the new document does not possess any pages yet.

➲ This command is especially handy if the same type of record has to be created again and again. You can copy one record and then use 'Paste record' several times to create as many of the same type of record as you wish.

31.4.4 Delete document from table (Ctrl+L)

Tutorial on Delete Document from Table (movie 2min) *5.4.1*

Use this command to remove a document together with its pages from the table. Before the record is deleted you must confirm your intention.

➲ Deletion of records is final! There is no function 'Undo delete'! Delete records only when you are absolutely sure that you do not need them anymore.

31.4.4.1 Deleting archived documents

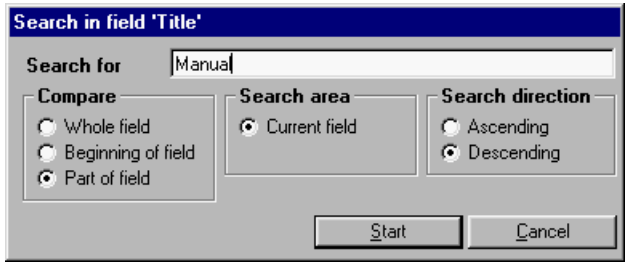
To protect users from inadvertently deleting parts of their archive it is not normally possible to delete an archived document. However, you can delete such a document after having deleted the check mark from the box 'archived' first. Let yourself be guided by the dialogue screens. Please note that documents and pages already transferred to external media are not affected by this, i.e. that they cannot be deleted this way.

31.5 Searching field content (Shift+F5)

Within the current selection of records you can search for data in a particular field. To search for a specific string of text you must move the cursor to the field you want to search, i.e. the field must be active (it does not play a role which document is active; all fields of the current selection of documents will be sifted through).

➲ Please note that menu 'Search' (see chapter 11.2) holds ready a much more comfortable possibility. The difference is that 'Searching field contents' uses by default the current selection while 'Searching in fields (documents)' by default creates a new selection.

After running the function (menu 'Edit'/Search field contents) the following dialog box appears:



The header tells you which field the search procedure pertains to. Please take note of the possibilities which are at your disposal in the frame **'Compare'**. There are three options:

'Part of field' (default): the search procedure will turn up every record that has a field in whose text the searched for term(s) occur(s).

'Beginning of field': Only those fields will be found which start with the search term.

'Whole field': Only those fields will be found which correspond 100% to the search string.

Furthermore, you can specify under **'Search direction'** if you want to search upward or downward in the list. The search starts as soon as you click the button **'Start'.** When an entry is found, you will be asked whether you want to proceed with your search.

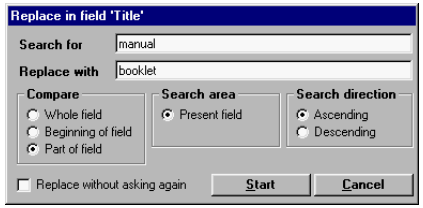
➲ Entering the value *NULL* as search term means that you search for all records which do not possess a value, i.e. which never were edited. Since such a search comprises the whole field you must choose the option 'Whole field' under 'Compare'.

Within **Yes/No fields**(e.g. field 'Marked') you can search for the Yes value by entering 'Y' or 'Yes', for the No value by entering 'N' or 'No'.

31.6 Replacing field content (Ctrl+F5)

Tutorial on Replacing Field Content (movie 2min) *5.3.1*

Archivista also offers you a function to automatically replace certain terms within your database. Please be aware of the fact that effected changes cannot be undone. This is especially important with regard to the option 'Replace without asking again'. The illustration below gives you the dialog box **'Replace'**. It is almost an exact correspondence to the dialog box 'Search' **'Search**. That is why only the differences are discussed here.



Before you replace a specific search term with another term you must enter the two text bits in the relevant text fields, 'Search for' and 'Replace with', respectively. Only then the procedure of replacing all terms can be started by clicking **'Start'.**

Normally, the replacement procedure occurs manually, i.e. you must confirm each replacement. By ticking the box 'Replace without asking again' the replacement of all terms occurs automatically.

Also when using the replace-function you can work with the expression ***NULL*.** You can use it for 'Search for' as well as for 'Replace with'. What is more, the expression ***ALL*** helps you to eradicate or completely replace a certain term from your database if you wish to do just that. Enter *ALL* under 'Search for' and confirm each single replacement (by another term or by *NULL*). Since such a procedure comprises the whole field you must choose this option within the frame 'Compare'.

➲ You cannot undo this function. All the replacements that have taken place are definitive, the process cannot be reversed. However, if you replace manually, i.e. without having ticked the option

'Replace without asking again', you will be asked for your confirmation before each definitive change.

31.7 Creating documents & scanning pages (Shift+F9)

When you are in table mode in table 'Archive' you can do several things at once with this function. You can a) create a document, b) change to 'Page view' and c) call up the mask 'Scanning with settings'. This helps you to scan your daily load of paper really fast.
This function depends on a scanner being connected to your computer. It must be turned on, too. If not, an error message will be issued.

31.8 Show pages

Use the following submenu items to navigate within the individual documents. The first four entries correspond to the buttons at the bottom of the detail view.

31.8.1 First

The first page of the document is activated.

31.8.2 Previous (Shift+F3)

The page preceding the current page is activated.

31.8.3 Next (Shift+F4)

The page following the current page is activated.

31.8.4 Last

The last page of the document is activated.

31.8.5 Go to... (Ctrl+G)

With this menu item you can go directly to the desired page in your document. When running this function you will be asked for the page number. Either enter the absolute value ('78') or enter the relative position (e.g. '+10' for 'ten pages further on' or '-12' for 'twelve pages back').

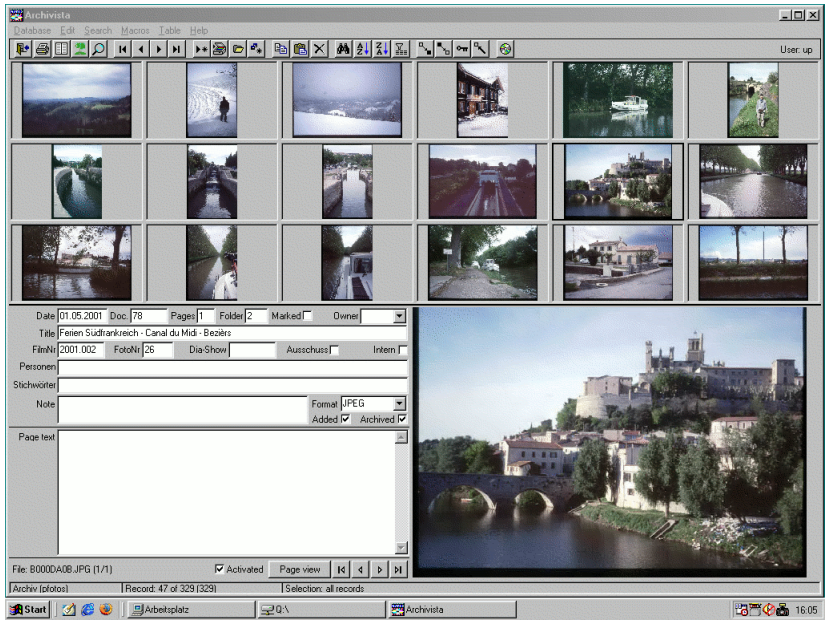
31.8.6 File number

This is a menu item which you do not ordinarily use. It is much more elegant and comfortable to go to the desired page by accessing the document first and then navigating to the desired page!
On the basis of the document number, the page itself and the resolution Archivista calculates a **file number** for each individual page. This number is at the same time the **file name** of the page. It always consists of eight letters / digits. In the following a short explanation:

1st letter: A=high resolution, B=screen copy resolution
2nd to 6th letter / digit: file number (encoded, taking a value from 0000B to ZZZZZ)
7th and 8th letter / digit: page number (encoded, taking a value from 0B to ZZ)

Note: In table 'Archive' a page will only be displayed in high resolution if no screen copy is found. Generally only the 'B'-page will be shown in cases where you enter a file number starting with an 'A'. In 'Page view', however, what is shown depends on what you enter. If you enter an 'A'-page, the 'A'-page will be fetched from the external medium and shown on screen. (This requires that you insert the relevant CD in your CD-drive or that you connect the corresponding external medium to your computer.)

31.9 Picture mode (Shift+Ctrl+F9)



The command **'Picture mode'** turns the main view into a view where the upper part of the screen shows preview images (instead of the table). Each document is represented by its first page/image. Typically this makes sense for photo collections.

This submenu can also be run by clicking the button .

31.10 Format (memo fields)

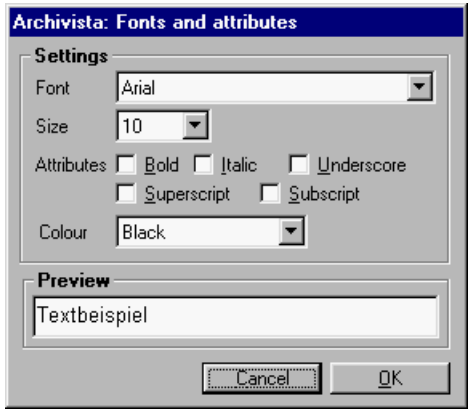
Tutorial on Format (movie 2min) *5.2.2*

Menu item **'Format'** can only be used within memo fields / RTF fields (Rich-Text-Format). These are fields in which font formatting (such as putting chars in **bold** or underlined type) is possible.

In Archivista 2009/I the following fields are RTF fields: field **'Note'** in tables 'Archive', 'Notes' and 'Literature' as well as the field **'Remarks'** in table 'Addresses'.

The format presently active will be ticked. See menu items under 'Format'.

All formatting possibilities can be specified under menu item **'Font formatting'** (see illustration below).



Note: We refrain from explaining this menu in more detail. It should be self-explanatory. Please be aware, though, that this menu can only be accessed when you are in an active RTF field (which is here always a memo field)!

32 Menu 'Search' and further menus

Tutorial on Menu 'Search' (movie 2min) *5.3.1*

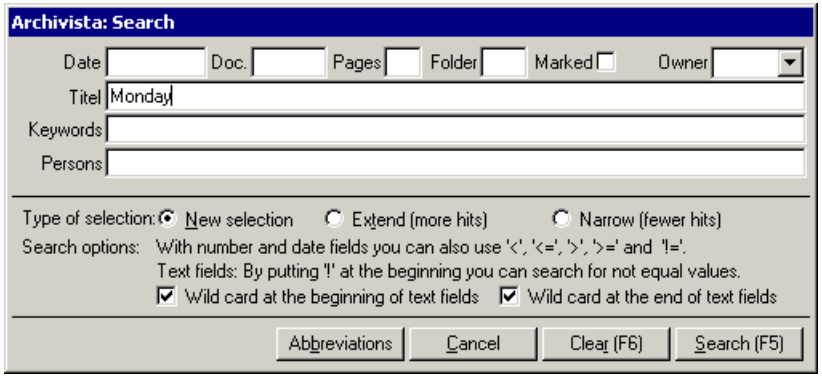
When you look for documents/records this is the menu to use. The first submenu 'Documents (fields)' serves to search for values in fields, the second 'Pages (fulltext)' enables you to search for words or strings of words in the pages (whose ASCII text was extracted).

The next two submenus help you to sort your data fast and comfortably in ascending or descending order. The other items are predefined selections or filters. They help you to make quick searches without having to use the search form. All records fulfilling certain requirements are shown in the list and can be edited. The other (not displayed) records remain somewhere in the background.

32.1 Documents (fields)...(F5)

Tutorial on Documents (movie 2min) *5.3.1*

Running this menu item is the simplest possibility to look for specific documents. You will see the following:



It is basically the same field mask that appears in the bottom left corner of the main view. In contrast to the main view where you change field values this mask serves to make a selection: simply enter the word or words for which you want to search.

➲ We recommend that you memorize function key 'F5' because with it you can call up the search mask without having to use the mouse. You can use 'F5' too to run the search after having filled in the mask.

Please take note of the buttons 'Clear (F6)' and 'Abbreviations.' With **'Clear (F6)'** you can delete previous entries and with **'Abbreviations'** you can show the abbreviations list for comfortable input of new search strings.

➲ The search mask possesses a double function: it mainly serves as search tool, however, it can be used for editing entries in table 'Field list', i.e. entries in pull-down menu fields. See section 36.6.7.1.

32.1.1 Type of selection

The default value is **'New selection'.** With this option the search takes into account all documents. Often you may want to build on your previous search. For example, you want to retrieve all documents which show the values 'moonstone' or 'topaz' in the title. Alternatively, you want all documents which have both 'moonstone' and 'topaz' in the title.

This you achieve with the options **'Enlarge (more hits)'** and **'Narrow (fewer hits)'**. They are no difficult options: we recommend that you create a few test documents for practising purposes. For example, create a number of documents and use expressions like 'moonstone,' 'moonstone-related articles,' 'gemstone,' 'topaz,' and 'precious stones' in the titles. Then, play around with the search options by building on your first search and entering the above terms in turn.

32.1.2 Search options

By means of the search options you can refine your search. Use '<', '<=', '>', '>=' and '!=' in front of a number or date value (only possible with the corresponding field types) to narrow down your search. Let us make a few examples. Create a few documents with the dates '01/01/2004,' '04/20/2004,' '07/10/2004,' and '01/01/05.' Now, run a selection with '<01/01/2004': none of the four documents should appear. When using '<=01/01/2004' you should see the first document in the hitlist and when using '>07/10/2004' the last two documents should appear. The formula '>=01/01/2005' returns all four documents while '!=01/01/2005' excludes the last document. When working with text fields only '!' for negation is at your disposal, i.e. by entering '!moonstone' the documents resulting from your search do not carry the word 'moonstone' in their titles.

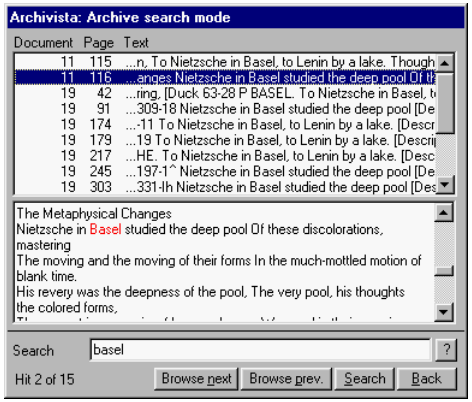
For text fields there are also the **wild cards.** The text comparison can thus be extended. 'Wild card at the beginning of text fields' means that a search with 'stone' also returns documents with 'moonstone' or 'gemstone' in the title. And 'wildcard at the end of text fields' would mean that 'moonstone' as search word would also return the document 'moonstone-related articles,' for example.

➲ As a rule, working with the wild card options is advisable. It pays to turn them off, though, when you get too many (wrong) hits. In such a case you can refine your search by unticking these options.

32.2 Pages (fulltext)... (Ctrl+F9)

Tutorial on Pages(fulltext) (movie 2min) *5.3.2*

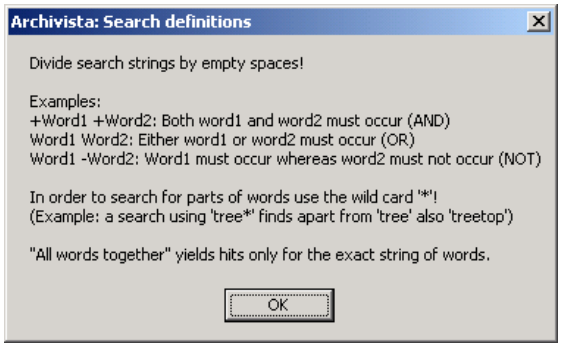
You can run this function both in the table mode or in 'Page view'! The 'Archive search mode' enables you to do a full text retrieval on all pages. Please note that the 'Archive search mode' works only for those pages previously treated with the OCR function. You find this command in **menu 'Search'** (or in menu 'Archive' when in 'Page view'). Also the shortcut **'Ctrl+F9'** or the **icon with the magnifying glass** take you to the 'Archive search mode'.


Archivista: Archive search mode
Document Page Text
11 115 ...n, To Nietzsche in Basel, to Lenin by a lake. Though
11 116 ...anges Nietzsche in Basel studied the deep pool Of th
19 42 ...ring, [Duck 63-28 P BASEL. To Nietzsche in Basel, t
19 91 ...309-18 Nietzsche in Basel studied the deep pool [De
19 174 ...-11 To Nietzsche in Basel, to Lenin by a lake. [Descr
19 179 ...19 To Nietzsche in Basel, to Lenin by a lake. [Descri
19 217 ...HE. To Nietzsche in Basel, to Lenin by a lake. [Desc
19 245 ...197-1^ Nietzsche in Basel studied the deep pool [De
19 303 ...331-Ih Nietzsche in Basel studied the deep pool [Des
The Metaphysical Changes
Nietzsche in Basel studied the deep pool Of these discolorations, mastering
The moving and the moving of their forms In the much-mottled motion of blank time.
His revery was the deepness of the pool, The very pool, his thoughts the colored forms,
Search
basel
?
Hit 2 of 15
Browse next
Browse prev.
Search
Back

To search for words in the pages of your archive you enter them followed by a space. Click the button **'Search'** to start the process. After some time you will get the search result in a list (only if anything is found). When you browse through the list now you will notice that the corresponding document and page are shown at the same time.

A maximum of 30 entries is displayed in the list. The buttons **'Browse next'** and **'Browse previous'** serve to move forwards and backwards. When you press the button **'?'** you get information on how to do combined searches.

32.2.1 MySQL search definitions


Archivista: Search definitions
Divide search strings by empty spaces!
Examples:
+Word1 +Word2: Both word1 and word2 must occur (AND)
Word1 Word2: Either word1 or word2 must occur (OR)
Word1 -Word2: Word1 must occur whereas word2 must not occur (NOT)
In order to search for parts of words use the wild card '*'!
(Example: a search using 'tree*' finds apart from 'tree' also 'treetop')
"All words together" yields hits only for the exact string of words.
OK

If you work with the indexing function of MySQL, there are the following options at your disposal (regarding fulltext retrieval): The operator '+' means that all the words entered must be present. If you enter **'+drink +alcohol'** you will find all pages containing both 'drink' and 'alcohol.' By entering **'drink alcohol'** you will find all pages containing either 'drink' or 'alcohol.' The search **'drink -alcohol'** returns all pages which contain the word 'drink' but do not contain the word 'alcohol.' Wild card search by means of the asterisk is also a possibility. When you enter **'drink*',** pages not only with 'drink' but also with 'drinks' or 'drinking' will be returned, however, pages with 'drunk' will not be selected (you would have to enter 'drink* drunk*'). Wild card search with MySQL is restricted to word endings for the time being. However, the search for an entire sentence is possible: **'We want drinks'** will return only pages which contain the phrase as it stands.

Note that MySQL indexes only words consisting of at least four letters. This is a default value that the client application Archivista 2009/Itakes over from the MySQL server. If you would like to change the settings of the fulltext index, you must make these changes directly in the MySQL server. More information about fulltext indexing you find in the MySQL manual.

➲ Let us disclose at this juncture only that you can change the minimum word length to 2 chars by going to the 'my.ini' start file and making the following alteration in a line following `'[mysqld]'`: 'ft_min_word_len=2' After such a change the MySQL server must be started afresh and the Archivista index must be rebuilt by running the function 'Create new index' in menu 'Database,' 'Parameters,' 'Archiving & index.'

32.2.2 Hitlist

The fulltext retrieval facility of Archivista allows you comfortable access to the pages you look for. However, the larger an archive the more hits will be returned by a fulltext search. That is the reason why the Archivista 2009/Ifulltext search works always in conjunction with the current selection of documents. I.e. we recommend that you narrow down the pages to be fulltext searched by doing a selection based on field content first.

In the final analysis a fulltext search which returns 34'234 hits does not help you much: you would not want to browse through all these pages. That is one reason why Archivista 2009/Ishows you by default a maximum of 1'000 hits per search.

32.2.2.1 Influencing the number of hits

Should you aim for an archive that is solely based on fulltext (something we expressly do not recommend), you might want to influence the maximum number of hits. This can be done in Archivista 2009/I, even if we have to warn you that increasing this number can severely affect performance of larger archives, especially in the network. In order to change this setting you need the console 'mysql' or a MySQL frontend. As a last resort you may use an ODBC connection.

➲ Please bear in mind that you change the values in question at your own risk. Certainly, no free support on this question could

be given. We recommend that you use the default values if you do not follow the description below at all.

For those who remain undeterred and are still willing to make these changes let us take a look at table 'parameter.' You will find two columns: **'Name'** and **'Inhalt'.** In column 'Name' you find two entries **'SearchHits'** and **'SearchMax'.** The default value vor 'SearchHits' is '9' and the one for 'SearchMax' is '1000.' Under 'SearchHits' you can enter any number between '9' and '50' in order to specify the number of hits appearing in the fulltext retrieval search window before you have to use the 'Browse' buttons. Under 'SearchMax' you can determine the maximum number of hits displayed after a fulltext search. You can use values between '100' and '20000.' Changes will only take effect from the next start-up of an Archivista client.

32.3 Sort (ascending) (Ctrl+F6)

Left-click a field or navigate to the desired field by pressing the tab key. Run this function and all the records in the current selection are sorted in an **ascending** manner.

➲ You can also trigger an ascending sort by going to the main view and clicking the column title of the field in question.

32.4 Sort (descending) (Ctrl+F7)

Left-click a field or navigate to the desired field by pressing the tab key. Run this function and all the records in the current selection are sorted in a **descending** manner.

➲ You can also trigger a descending sort by going to the main view, **keeping the shift key pressed** and clicking the column title of the field in question.

32.5 All Records (F6)

When you start up Archivista the first found records will be displayed in each table. This 'original state' you can get back to at any time by the function **'All Records'**. You simply need to press **'F6'.**

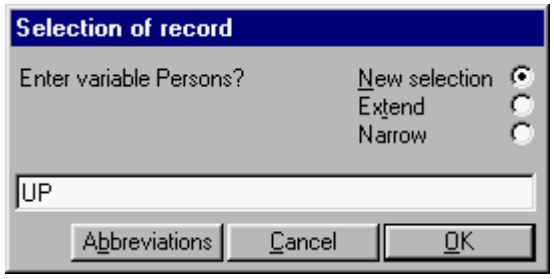
➲ Please note that not all records of a query may be shown. The status bar at the bottom of the window gives you the number of the documents found beside the figure in brackets, which gives the maximum number of records displayed for the current user. More information on this is to be found in 35.2.1.3.

32.6 Working with selection criteria

There are times when you do not want to work with all existing records. In these cases the defined criteria / filters in menu 'Selection' may be helpful. They are listed under the default selection **'All records'.**

Click the desired filter. Subsequently the selection process gets started and may take a few seconds.

Please note that there are filters or selection criteria which possess variables. After running such a filter you will be asked for further information necessary to do the selection. A variable in the field 'Persons' means that you have to specify which person you want to include in your search criteria. The following screen appears:



Enter a person. Press 'OK' and the specified selection will be run.

➲ You can also use wildcard symbols (asterisks), for example, if you only know the beginning of the person's name. 'Buch*', for example, will find all persons named 'Buchan' as well as all persons called 'Buchanan' or 'Buchman'. Wildcard symbols can be used both at the beginning or the end of a search term.

32.6.1 Extending or narrowing down of selections

Each time you run a selection (Menu 'Selection') for which variables must be entered (normally these are the definitions with three question marks) you can decide, if you want to extend or narrow down your search. By ticking **'Extend'** you get a Boolean OR, i.e. records will be added to your previous selection, and by ticking **'Narrow'** you effect a Boolean AND, i.e. only those records will be shown for which this and the previous selection criteria apply.

32.7 Menu 'Macros'

The menu 'Macros' is intended for your own macros and can contain up to 9 user defined command sequences per main table. You will find more information on the macros in Archivista 2009/I in chapter 35.8.

32.8 Menu 'Table'

The menu 'Table' allows you to go to the different tables in Archivista 2009/I. One must differentiate between the **main tables** 'Addresses', 'Notes', 'Archive' and 'Literature' and the **auxiliary tables** 'Abbreviations', 'Field lists,' 'Zip codes (Addresses)' and 'Rubrics (Literature)'.

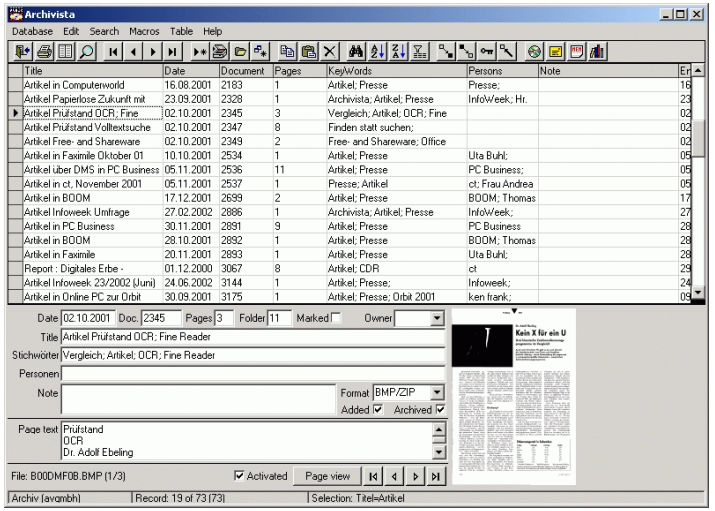
In contrast to the main tables the auxiliary tables possess only limited functionality. There are no shortcuts and no icons in the tool bar.

32.8.1 Activate list/form (F11)

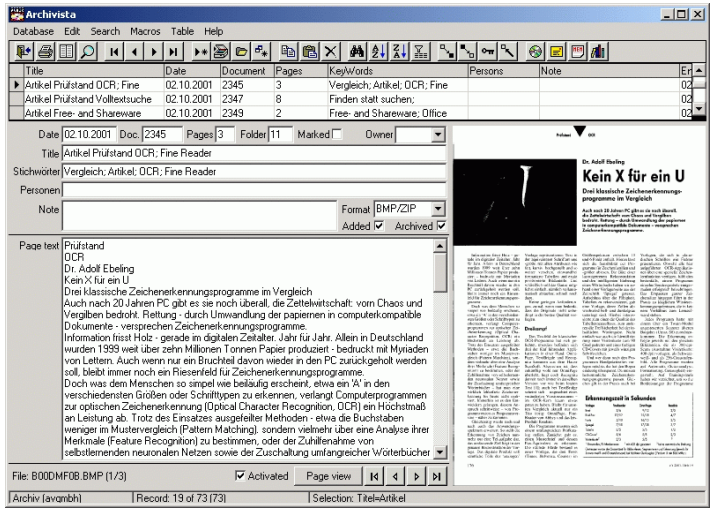
Use this function to switch from a field in detail view to the corresponding field of the corresponding document in the list and back again. Navigation through your documents is very much simplified.

32.8.2 Enlarge/diminish list (Ctrl + F11)

In standard view the active page is normally displayed as a very small picture in order to list as many documents as possible. Sometimes it makes sense though to display only a short list and have a larger picture instead. See the following two illustrations:


Archivista
Database Edit Search Macros Table Help
Title Date Document Pages KeyWords Persons Note
Artikel in Computerworld 16.08.2001 2183 1 Artikel; Presse Presse;
Artikel Papierlose Zukunft mit 23.09.2001 2328 1 Archivista; Artikel; Presse InfoWeek; Hr.
Artikel Prüfstand OCR; Fine 02.10.2001 2345 3 Vergleich; Artikel; OCR; Fine
Artikel Prüfstand Volltextsuche 02.10.2001 2347 8 Finden statt suchen;
Artikel Free- and Shareware 02.10.2001 2349 2 Free- and Shareware; Office
Artikel in Faximile Oktober 01 10.10.2001 2534 1 Artikel; Presse Uta Buhl;
Artikel über DMS in PC Business 05.11.2001 2536 11 Artikel; Presse PC Business;
Artikel in ct, November 2001 05.11.2001 2537 1 Presse; Artikel ct; Frau Andrea
Artikel in BOOM 17.12.2001 2699 2 Artikel; Presse BOOM; Thomas
Artikel Infoweek Umfrage 27.02.2002 2886 1 Archivista; Artikel; Presse InfoWeek;
Artikel in PC Business 30.11.2001 2891 9 Artikel; Presse PC Business
Artikel in BOOM 28.10.2001 2892 1 Artikel; Presse BOOM; Thomas
Artikel in Faximile 20.11.2001 2893 1 Artikel; Presse Uta Buhl;
Report : Digitales Erbe - 01.12.2000 3067 8 Artikel; CDR ct
Artikel Infoweek 23/2002 (Juni) 24.06.2002 3144 1 Artikel; Presse; Infoweek;
Artikel in Online PC zur Orbit 30.09.2001 3175 1 Artikel; Presse; Orbit 2001 ken frank;
Date 02.10.2001 Doc. 2345 Pages 3 Folder 11 Marked Owner
Title Artikel Prüfstand OCR; Fine Reader
Stichwörter Vergleich; Artikel; OCR; Fine Reader
Personen
Note
Format BMP/ZIP
Added Archived
Page text Prüfstand OCR Dr. Adolf Ebeling
File: B00DMF0B.BMP (1/3)
Activated
Page view
Archiv (avqmbh)
Record: 19 of 73 (73)
Selection: Titel=Artikel

After starting the program you get the above view: small picture of page, long list.


Archivista
Database Edit Search Macros Table Help
Title Date Document Pages KeyWords Persons Note
Artikel Prüfstand OCR; Fine 02.10.2001 2345 3 Vergleich; Artikel; OCR; Fine
Artikel Prüfstand Volltextsuche 02.10.2001 2347 8 Finden statt suchen;
Artikel Free- and Shareware 02.10.2001 2349 2 Free- and Shareware; Office
Date 02.10.2001 Doc. 2345 Pages 3 Folder 11 Marked Owner
Title Artikel Prüfstand OCR; Fine Reader
Stichwörter Vergleich; Artikel; OCR; Fine Reader
Personen
Note
Format BMP/ZIP
Added Archived
Page text
Prüfstand
OCR
Dr. Adolf Ebeling
Kein X für ein U
Drei klassische Zeichenerkennungsprogramme im Vergleich
Auch nach 20 Jahren PC gibt es sie noch überall, die Zettelwirtschaft: vom Chaos und Vergilben bedroht. Rettung - durch Umwandlung der papiernen in computerkompatible Dokumente - versprechen Zeichenerkennungsprogramme.
Information frisst Holz - gerade im digitalen Zeitalter. Jahr für Jahr. Allein in Deutschland wurden 1999 weit über zehn Millionen Tonnen Papier produziert - bedruckt mit Myriaden von Lettern. Auch wenn nur ein Bruchteil davon wieder in den PC zurückgeholt werden soll, bleibt immer noch ein Riesenfeld für Zeichenerkennungsprogramme.
Doch was dem Menschen so simpel wie beiläufig erscheint, etwa ein 'A' in den verschiedensten Größen oder Schrifttypen zu erkennen, verlangt Computerprogrammen zur optischen Zeichenerkennung (Optical Character Recognition, OCR) ein Höchstmaß an Leistung ab. Trotz des Einsatzes ausgefeilter Methoden - etwa die Buchstaben weniger im Mustervergleich (Pattern Matching), sondern vielmehr über eine Analyse ihrer Merkmale (Feature Recognition) zu bestimmen, oder der Zuhilfenahme von selbstlernenden neuronalen Netzen sowie der Zuschaltung umfangreicher Wörterbücher
File: B00DMF0B.BMP (1/3)
Activated
Page view
Archiv (avgmbh)
Record: 19 of 73 (73)
Selection: Titel=Artikel
Kein X für ein U
Drei klassische Zeichenerkennungs-programme im Vergleich

After running menu item 'Enlarge/diminish list (Ctrl+F11) the list is shortened to make room for a bigger picture.

➲ This function is available only within table 'Archive'. Please note too that this function is akin to the function 'Enlarge page' described earlier on.

32.9 Menu 'Help (F1)'

By pressing the function key **'F1'** you will get instant access to the Help system of Archivista 2009/I – from wherever you happen to be in the program. Browse through the help system as you would be browsing through this manual.

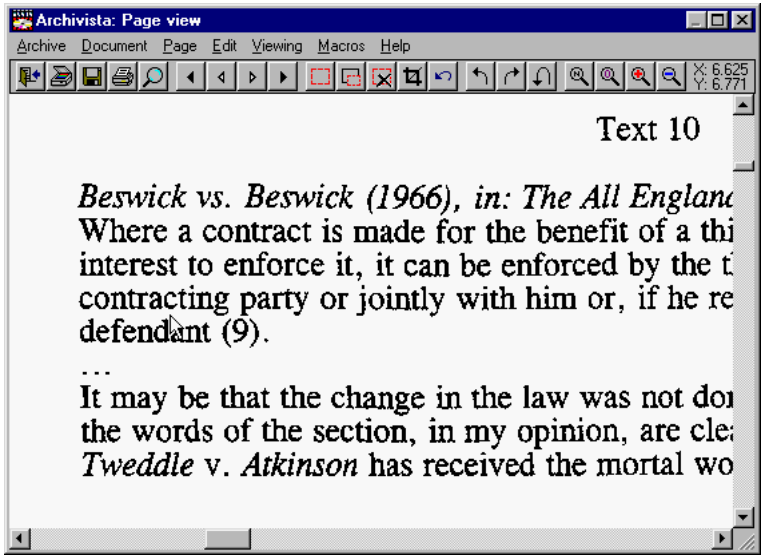
➲ Under 'Last news' you will find information which has not made its way into the manual yet.

33 Page View

Tutorial on Page View (movie 2min) *5.2.1*

33.1 What is 'Page view'?

With 'Page view' you can view individual pages as image files, you can turn them, modify, print or export them and, of course, you can add new pages. You can call up 'Page view' by going to **menu 'Database' / item 'Page view'** or you press **function key 'F9'.** You can also use 'F9' to move from 'Page view' back to the normal table mode. Also a **double click with your left mouse key on the page image** enables you to move between 'Page view' and the table mode.


Archivista: Page view
Archive Document Page Edit Viewing Macros Help
X: 6.625
Y: 6.771
Text 10
Beswick vs. Beswick (1966), in: The All Englanc
Where a contract is made for the benefit of a thi
interest to enforce it, it can be enforced by the t
contracting party or jointly with him or, if he re
defendant (9).
...
It may be that the change in the law was not dor
the words of the section, in my opinion, are cle
Tweddle v. Atkinson has received the mortal wo

'Page view' shows one page at a time, normally. Depending on whether the whole page can be displayed or not, below and on the right scrolling bars appear. With these you can view the rest of the page. You can also move through the page – and perhaps more elegantly – by means of the **'Arrow keys'** (on the right of your keyboard, close to the right shift key).

A description of the menus you will find below. Again, the **toolbar** has the most important commands ready. What is more, to the right of the toolbar the **current position of the mouse pointer** is indicated ('X:' AND 'Y:'). The values given here are based on the settings in menu 'Viewing'(see chapter 33.6.13).

If your monitor possesses a resolution higher than 640x480 pixels you get additional information on the page displayed, namely: the folder, the document, the current page, the number of pages in the document, the file name as well as the title of the document.

33.1.1 Menu structure in 'Page view'

Global functions, i.e. functions that have an effect on all documents, are to be found in menu 'Archive'. Functions referring to the document are in menu 'Document' and functions pertaining to whole pages are accessible via menu 'Page'. Menu 'Edit' helps you to work on individual pages and menu 'Viewing' serves to display your pages in different ways.

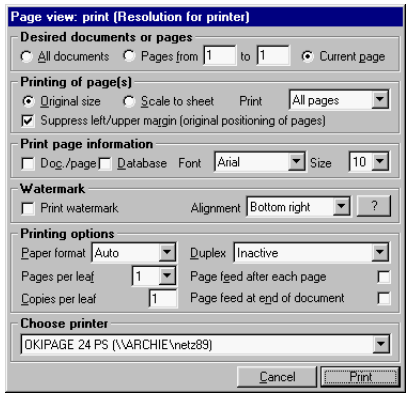
33.2 Menu 'Archive'

33.2.1 Print... (Ctrl+P)

Tutorial on Print (movie 2min) *5.4.3* **Tutorial on Print (movie 2min)** *5.4.4*

Note: When printing in 'Page view' only the pages are printed. If you want to print other document information you must go

to **table mode** where you run the menu item 'Print' in menu 'Database'.



Before printing, you need to specify the number of pages you wish to print. **In 'Desired document or pages',** you determine whether you want to print all pages, a specific number of pages or only the current page.

'Printing of page(s)' lets you choose whether the page will be printed in its original size (in which case not all of the page's contents may fit the printout format) or if it is to be scaled before being printed (i.e. that it fits the page). As most printers have set minimum margins that cannot be printed over, **'Suppress left/upper margin...'** allows you to override these margins. Under 'Printing' you may also choose between printing all pages, even pages only or odd pages only.

Activating **'Print page information'** will print the name of the document/page. The desired font and size can be specified. The default is 'Arial' 10 point.

The first choice under **Printing options'** is 'Paper size'; you can choose between 'Automatic', 'Portrait' or 'Landscape'. 'Automatic' means that the page orientation will be set according to the bitmap dimensions i.e. whether it will fit the page better as portrait or landscape. 'Duplex' lets you choose between single-sided or double-sided printing (depending on printer support). Under 'Pages per leaf' you can specify how many pages are fitted on each sheet. Up to 16 pages can be printed on the same sheet and you can also determine whether you want to have the first page of a new document on a new sheet with 'Page feed at end of document'. The option 'Page feed after each page' closes the print job after each page and creates a new job for the following page. On a network, this may speed up the printing process.

➲ Important: When printing Archivista always uses the page quality that is activated. If you want to print a page in the original high quality (and not in screen quality), then you must load the original page (function key 'F7').

33.2.2 Archive search mode (Ctrl+F9)

By running 'archive search mode...' you start a full text retrieval searching through all indexed pages of the archive. This function is described in depth in chapter 32.2.

33.2.3 Count pages in archive (Ctrl+J)

When you run this menu item the selected pages (=the pages of the documents in the selection in table mode) are counted and the result is shown.

33.2.4 Choose scanner...

If more than one scanner is connected to the system, you can select the one you want to use by this menu item.

33.2.5 Back to Archivista (F9)

With this menu item you quit 'Page view' and get back to the table mode, the main view of table 'Archive' in Archivista.

33.3 Menu 'Document'

33.3.1 Basic functions

Use the first four items of this menu to browse among the individual documents.

33.3.2 Sort...

Tutorial on Sort (movie 2min) *5.4.2*

This functionality is highly useful if the feeder of your scanner handles simplex mode 'only'. Scanning originals which are printed on both sides can be done in no time this way. If your scanner does not have a feeder (ADF) at all or if it has a duplex feeder, the sorting functions described below do not make any sense.

➲ Before running this menu item please make sure that you are familiar with the details of how it works. Otherwise you might inadvertently bring your pages into utter disorder. The following examples are based on the simplest case where there are eight pages involved. When there are more pages the page numbers change accordingly.

33.3.2.1 Normal

You scan all front sides of a batch of sheets first and then you do the reverse sides. The pages appear in the order **1,3,5,7, 8,6,4,2** now and need sorting so that in the end they appear in the 'natural' order 1,2,3,4, 5,6,7,8.

➲ Only those pages get sorted which lie behind the currently active page. This procedure has the considerable advantage of enabling you to alternately add and sort pages several times.

33.3.2.2 Booklets and Extended

These two options are more demanding than the one discussed above. We recommend that you practice with a small number of pages before your pages are irrevocably in disorder.

Booklets: The pages appear in the order **8,1,6,3, 4,5,2,7** and need sorting. This constellation may occur when you scan booklets with the staples taken out.

Extended: You may find this function useful if you want to scan papers which are the result of your previously copying a book in duplex mode on a copier. The pages appear in the order **1,2,5,6, 7,8,3,4** and need sorting.

If you want to sort your pages with the functions 'Booklets' or 'Extended', respectively, please take care to activate the option 'Divide page in the middle to make two pages' (menu 'Page' / 'Scanning with settings')).

➲ Here too only those pages get sorted which lie behind the currently active page. The total of the pages that you want to sort must be a figure that can be divided by four.

33.3.3 Remaining pages to new document (Ctrl+W)

You may need this option after **scanning with the document feeder.** After scanning a whole batch of sheets (they are automatically allocated to the currently active document) you can allocate pages to subsequent documents.

Example: You scan 9 pages which really belong to three documents. (In the end each of the document is supposed to contain 3 pages.) After scanning with the document feeder all pages 'belong' to document 1. Now you reallocate the pages that belong to document 2 and 3. Go to document 1, page 4. Run the command 'Allocate to next document'. Result: pages 4 to 9 are moved to document 2. Go to document 2, page 4. Run the command

'Allocate to next document'. The last three pages are moved to document 3.

33.3.4 Create new document (Ctrl+N)

With this menu item you create a new document while you are in 'Page view'.

33.3.5 Create document and scan ... (Shift+F9)

With this menu item you create a new document and call up the mask 'Add pages by scanning' in one step.

33.3.6 Mark pages for merge)

With this menu item you can 'earmark' all pages from a document for later merger with another document.

33.3.7 Merge marked pages with active document

With this menu item you can assemble the previously 'earmarked' pages with the currently active document.

33.4 Menu 'Page'

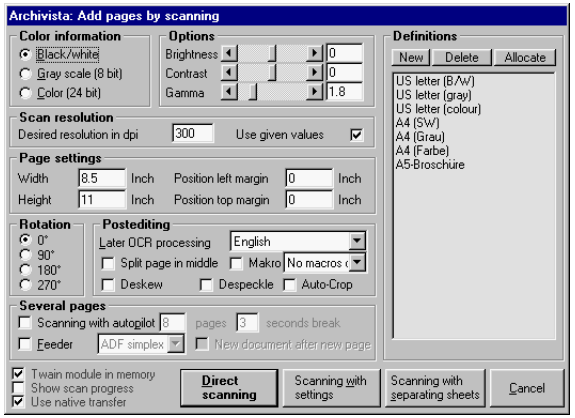
33.4.1 Basic functions

Use the first six items of this menu to comfortably browse through one document. These menu items are described in detail in chapter 31.8.

33.4.2 Scanning with settings (F2)

Tutorial on Scanning with settings (movie 2min) *5.2.1* **Tutorial on Scanning with settings (movie 2min)** *5.4.2*

This function serves to add pages to your document by means of a scanner. When you run it the following mask appears:



Please note: The function **'Scanning'** works only if a scanner is connected to your computer. The scanner must be Twain compatible.

33.4.2.1 Color information

Within this frame you determine the **color depth** of the scanning process, i.e. you specify if the page is read as a black and white page, as one with gray scales or as a color image. Black and white scans possess only black and white dots while images in gray scales contain different shades of black and white (and correspond to a black and white picture shot with a camera).

33.4.2.2 Options

Here you specify brightness, contrast and gamma value of the page you want to scan. If your pages are of satisfactory quality you do not normally need to change the settings. However, if your pages are of bad quality these options may do wonders!

33.4.2.3 Scan resolution

Under **'Scan resolution'** you specify the desired resolution when scanning. It goes without saying that the chosen value must be

supported by your scanner. Normally you would use 300dpi for black and white pages and 150dpi for pages in gray scales and color. At any rate, this is what you get when you tick **'Use given values'.**

The size of the page is specified under **'Page settings'.** It is a good habit to put the pages always in the same fashion on the pane of the scanner. The function **'Rotation'** serves to subsequently rotate landscape pages and put them on their feet.

33.4.2.4 Post-editing functions

Scans that consist of two pages side by side (as with the scanning of books) can be split into two separate pages by activating the **'Split page in middle'** function.

After scanning, you can automatically run a macro. You could, for example, brighten up the page automatically. In order to do this, you need to create the macro first (see tables 'Parameters' and 'Macros').

'Deskew' automatically straightens scans that are slightly askew (e.g. from scanning with a feeder). **'Despeckle'** removes dirt spots scanned from the sheet. **'Auto-Margin'** removes superfluous white borders.

➲ Note: The functions 'Deskew', 'Despeckle' and 'Auto-Margin' cannot fix serious scanning flaws. Also, when deskewing scanned pages containing plenty of visual data but very little text, it may be that they are not straightened up fully satisfactorily. This is because no software has yet been able to match the perfection of the human eye.

'Later OCR' is an option to determine if the page image should be turned into a text file later by means of an OCR software. If the material you scan is not text but images, you should choose 'Exclude' here to suspend OCR. Further information you get in chapter 30.7.

33.4.2.5 Scanning with autopilot

If your scanner does not have an ADF but you want to add several pages to the same document and all have the same settings, choose **'Scanning with autopilot.'** With this option turned on your scanner makes scans in regular intervals of a few seconds and all you need to do is to change the pages in between scans. You must enter the total number of scans that should be automatically made (1-640) and also the scanning interval must be specified.

33.4.2.6 Scanning with feeder (ADF)

If your scanner possesses a document feeder, do tick **'Feeder'** and choose **'ADF simplex'.** With this option on you can fill the document feeder, press 'Direct scanning' and walk away. Your scanner will faithfully work through the whole stack and it may be minutes before you have to attend to it again. Some (more expensive) scanners are able to scan both sides of a sheet at the same time. If you posses such a scanner, choose **'ADF duplex'.**

Your scanner and Archivista 2009/I will work perfectly together to add all the desired pages to your archive. You can cancel the process by pressing the **ESC key.** Please note that cancellation takes effect only after scanning of a page is completed.

The option**'Create new document with new page'** makes most sense when you work with a scanner that has a feeder. It allows you to scan whole stacks of paper and Archivista 2009/I allocates each new sheet to a new document.

33.4.2.7 Starting the scanning process

After entering all the desired settings you may start the **scanning process.** Press one of the following buttons to do that:

Direct scanning: the scanner is addressed directly and not via the scanning software dialog box supplied with the scanner.

Scanning with settings: the scanning software dialog box appears (this is not Archivista 2009/I but the scanning software belonging to the scanner).

The page is scanned. Afterwards you can **modify your page** (see menu 'Edit'), i.e. you can rotate it, remove flaws by cutting them out etc. What you cannot do is modification pixel by pixel.

➲ After adding the pages they are automatically stored. You can modify them until they are archived but you need to manually store these modifications (menu 'Archive' / 'Store current page').

33.4.2.8 Options for addressing Twain driver

Please take note of the option **'Twain module in memory'.** By ticking it you effect that the Twain driver is kept in memory after scanning. Apart from avoiding scanning problems with older drivers (16 bit drivers) this option makes the scanning process faster because the twain module does not have to be loaded before each scan. By choosing **'Show scan progress'** you see a bar on the screen during each scan which tells you roughly at what stage the scanning process is. By choosing the option **'Use native transfer'** the page will be scanned in the RAM and not in the hard disk buffer. It may well be that by choosing this option you will be able to scan faster, but not all scanners support it.

33.4.2.9 Definitions

Tutorial on Definitions (movie 2min) *5.4.2*

By means of the definitions you can take a snapshot of the settings that you use most often. Later you click a definition and the desired settings are active. In that way you save time and energy when you have to quickly change from a US letter page in black and white to a smallish piece of paper that needs to be scanned in grayscales.

To create a new definition change the settings according to your needs (but do not start the scanning process yet). Click the button 'New' (under 'Definitions') and enter a name for it. The new definition is ready and appears in the list as a new entry.

You call up an already **existing definition** by simply clicking the relevant entry with your left mouse key. Immediately the settings spring to the defined values. It goes without saying that you can **delete** definitions with the appropriate command. Also changing definitions is possible: call up a definition, change the values you want to change and click **'Allocate'**.

➲ Definitions are kept only during the current session. If you want to make them available to more than one user or for longer than the current session, they must be created with systems operator privileges. After quitting the program they are stored and at the disposal of all the users.

33.4.3 Direct scanning (Ctrl+F2)

You can start a scanning process without having to change settings beforehand. The previously entered settings are used.

33.4.4 Store current page (Ctrl+S)

After the pages are added they are automatically stored. From this point in time until the document is archived, you must manually store each page that you have altered. In these cases only a status message in the header tells you that the page is being stored (no confirmation box).

33.4.5 Delete page (Ctrl+L)

You can easily remove individual pages from a document. However, only up to the point in time when the document gets archived. You will get a message asking you to confirm the deletion.

➲ By pressing the Shift key you can delete several pages at a time.

33.4.6 Move current page... (Ctrl+M)

Normally, pages are added in the same order as they should appear in the archive and moving them should not be necessary. However, if a page scanned and added at a later stage has its rightful position early in a document you can move it by using this function.

➲ You can move the page backwards or forwards to any position you like.

33.4.7 Importing pages

With this function you comfortably import pages that already exist as bitmap files. The formats BMP, TIF, PNG and JPEG are supported.

33.4.7.1 Import from file... (Shift+F2)

With this menu item you import an individual bitmap file as a new page.

33.4.7.2 Import from clipboard (Ctrl+V)

A graphics page (also one previously created by Archivista itself) 'parked' on the clipboard may be added to any Archivista document this way.

33.4.7.3 Several pages from parameter file

With this menu item you can import several pages at once (with the help of a parameter file). All these pages are imported into the same document. The parameter file consists of a header and the names of the files to be imported:

```
*****Archivista-Parameterdatei*****
C:\Programme\Av5e\TEMP\E0000B0B.BMP
C:\Programme\Av5e\TEMP\E0000B0C.BMP
```

33.4.7.4 TIFF file (multipage) (Ctrl+I)

The distinctive mark of tiff multipage files is that they contain several pages in one graphics file. To import all pages of a tiff multipage file run this menu item.

33.4.8 Export

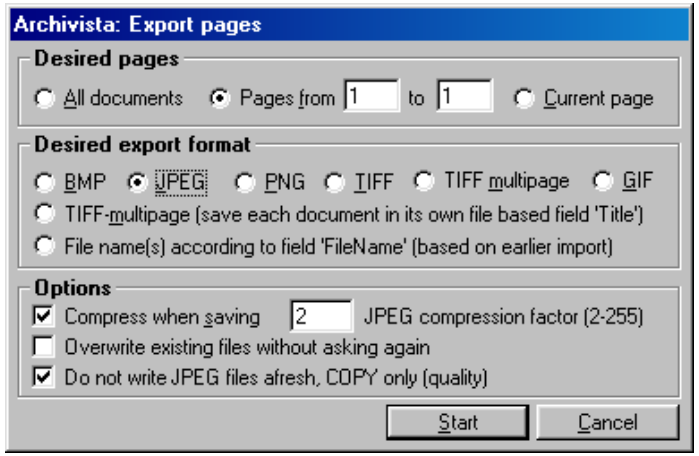
Use this function to export pages from Archivista to another application.

33.4.8.1 To file(s)

Use this menu item to export. The formats 'BMP', 'JPEG', 'PNG', 'TIFF', 'TIFF Multipage' and 'GIF' are at your disposal. The 'BMP' format is a Windows standard and can be read by all other Windows programs, you may need the other formats if you work with other operating systems (such as Macintosh).

Under **'Options'** you decide whether **TIFF** and **JPEG files** are filed in compressed form or not when exported. When exporting **JPEG** files a compression factor may be entered (2=lowest compression, 255=highest compression). In addition to that you may specify whether an alert screen should warn you against overwriting. And by ticking the third option you determine that for better quality JPEG files should be **copied** when exported (not written afresh).

After pressing 'Start' the pages are being exported. The process may take a while. Cancellation is possible, however, with the button 'Cancel'.



To export the specified pages click 'Start'. In the case of a Tiff multipage file you must name the graphics file yourself. When exporting several pages in the formats 'BMP', 'JPEG' or 'TIFF' a parameter file is created. The file names of the individual pages are composed of the letter 'E' and the ending of the filename previously used by Archivista. With the help of the parameter file pages exported in this way can be allocated to a different document, for example.

➲ The option 'TIFF-multipage (save each document in its own file based on the field 'Title')' is only useful if you have chosen to export 'All documents'. In this case, every document will be exported as a separate file. The file name is derived from the document's 'Title' field. Please note that before exporting you will need to define a parameter file (suffix '.avp') which after exporting will contain a report of how many pages were exported from each document.

33.4.8.2 To clipboard... (Ctrl+C)

You can also copy the active page (=the one currently displayed on the screen) to the clipboard in order to make it available to another program. Please note, that this operation may take some time depending on the RAM available on your computer.

33.4.8.3 Page to clipboard

In RTF format, the current page can be exported directly to the clipboard.

33.4.9 OCR settings current page

Use the items of this submenu to determine how your active page should be treated with the OCR software. More detailed information is to be found in chapter 30.7.'

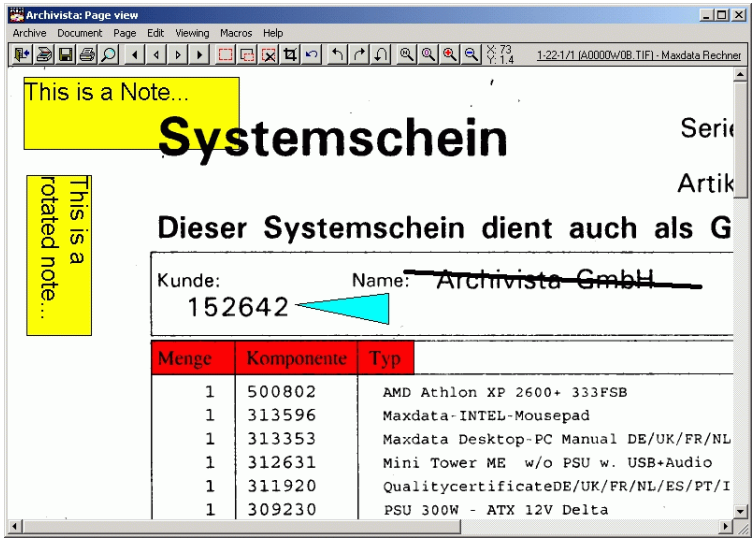
33.4.10 Notes

Each additional item of information for pages input into the system should preferably be recorded in a (fixed!) database field, rather than in an unstructured way, i.e. the equivalent of having lots of small pieces of paper lying around. However in certain cases (e.g. for printouts or forwarding a digital document), 'stick-on' notes can be useful.

To activate the notes, go to **Page view** in the **'Page' menu** and from the **'Notes' sub-menu** select **'Add'.** If the database does not yet contain any notes, you will be asked if this functionality is to be activated now. If you **click on the 'Yes'** button, the database will be adjusted accordingly. After a **Restart** you will be able to add the 'Notes' menu item.

Each time you 'Add' a note, a frame-type note will appear in the top left-hand corner. Click on this, pressing the **right-hand mouse key.** This takes you directly to the **'Notes' sub-menu;** you can now set the parameters. You can also get to notes (when notes are not activated) by holding the Shift key down and then pressing the right-hand mouse key.

Under **'Type'**you can specify whether the note is to be displayed as a **Line, Triangle, Frame or Text box.** A note is normally activated by clicking on it. Alternatively, pressing 'Select' will also take you to the editing function. A number of examples are given below:



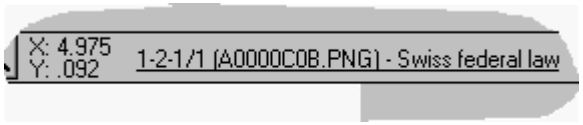
Go to menu item **'Duplicate'** to make copies of existing notes (the copy will be superposed on the former note). To delete a note, select **'Delete',** and the edited notes will be **saved automatically.**

Additional parameters are: **Position, Colors** and **Line width.** These menu items should be self-explanatory. However note that you can create **transparent notes** by going to 'Colors' and sub-menu item 'Translucence'.

If you use the 'Type' menu item to open a **text box,** you will first see the text 'Archivista'. If you **double-click** on this note, you will be able to edit the text directly. Text boxes can also be rotated in 90-degree steps.

Notes are administered via the database, so you can re-edit them at any time. To physically attach objects to the page (e.g. for an export operation), use **'Store with image'.** The notes will then be merged with the current page (for the final save, press Ctrl+S).

You can use **'Deactivate'** to specify whether the notes are to be displayed or hidden. If you have selected 'Deactivate', the pages will be displayed more quickly. This also means that the notes will not appear on page print-outs. If 'Deactivate' is not ticked, the notes will automatically be printed. You can see whether or not a page contains notes by looking at the page view **title line.** An **underlined entry** means that the page in question contains **notes.**



Final point on 'Notes': The notes are provided solely for 'Page view'. However, it makes no difference whether you have requested the high-resolution A copy or smaller B copy.

33.4.11 File info of current page (Ctrl+A)

This menu item gives you information on resolution and size of the currently active page.

33.5 Menu 'Edit'

This menu gives you a variety of possibilities to modify added pages. If we were to describe all functions in detail the result would be a boring listing. Instead, we recommend you to try them out one by one and to test for yourself what these commands can do!

➲ Archivista is a document management system and no image processing software! That is the reason why you cannot edit the added pages pixel by pixel. In particular you cannot add content

with paint brushes and the like. But you can modify the entire page by rotating or brightening it, for example.

33.5.1 Area (size)

Tutorial on Area (movie 2min) *5.4.1*

Use the upper eight commands in the **'Edit'** menu to modify an area of the currently active page or to change its size.

➲ Important: the mouse is absolutely essential for the next five menu items!

33.5.1.1 Mark (Ctrl+Ins)

When you run this function the mouse pointer changes (the arrow turns into a cross). Move the mouse to the upper left corner of the area you want to mark. Press the left mouse button, hold it down and move the mouse button to another location: a frame appears. As soon as you release the left mouse button the frame is copied onto the clipboard.

33.5.1.2 Paste from clipboard (Shift+Ins)

When you run this function the mouse pointer changes into an arrow pointing upwards. Move the mouse to the location where you want to insert the frame previously marked. Left-click your mouse and the frame from the clipboard is inserted.

33.5.1.3 Mark and delete (Shift+Del)

The procedure is identical to the one described under 'Mark' (Ctrl.+Ins). However the effect is totally different in that the marked area is not copied but deleted.

33.5.1.4 Cut page in chosen area (Alt+Backspace)

When running this function the mouse pointer changes into a cross. Move the mouse to the desired location and press the left mouse button. Hold it down and move to another position: a frame appears. Release the mouse button and the page is reduced to the size of the frame.

33.5.1.5 Cut page bottom/right (Ctrl+U)

The mouse pointer changes into an arrow pointing upwards. Move your mouse to the location which you want to be the future bottom right corner of your page. Left-click and the page is cut below and on the right and reduced to the desired size.

33.5.1.6 Modify page width...

After entering the new page width the page is cut on the right side. At this point in time only a reduction of the width is possible.

33.5.1.7 Modify page height...

After entering the new value for page height the page is cut at the bottom. At this point in time only a reduction of the height is possible.

33.5.1.8 Divide page in two (Ctrl+T)

Tutorial on Divide page in two (movie 2min) *5.4.1*

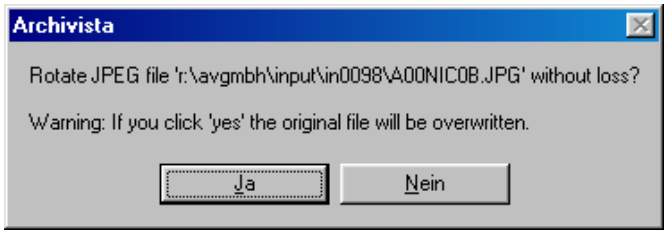
This menu item divides a page vertically in the middle. The left side stays with the old page, for the right side a new page is created.

➲ Only the last page can be divided! It goes without saying that archived pages cannot be divided anymore.

33.5.2 Rotating pages

With this function you can rotate your pages to your heart's content. You can use the predefined values (90°, 180° and 270°) or enter any value you like (a figure between 0 and 359 degrees).

When you rotate a **JPG image** with predefined values the following dialog box will pop up:


Archivista
Rotate JPEG file 'r:\avgmbh\input\in0098\A00NIC0B.JPG' without loss?
Warning: If you click 'yes' the original file will be overwritten.
Ja
Nein

In order to rotate the image without loss Archivista 2009/I uses a special subroutine which replaces the original image straightaway. If you are sure that you want to save the image in its rotated state, then this dialog box must be confirmed.

33.5.3 Modifying pages

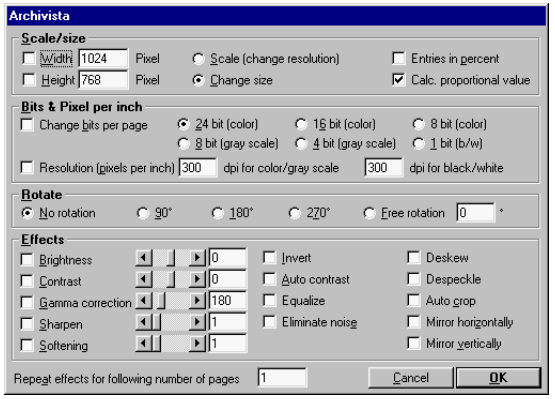
These menu items either alter the resolution of the page or change its size.

33.5.4 Effects

Under **'effects'** a variety of functions are brought together which serve to modify the content of a page (but not its size or resolution). You will find here the most widely used functions such as inversion, brightness, contrast, saturation, gamma correction, sharpening and softening contours.

33.5.5 Using the dialog box to modify editing parameters (Ctrl+H)

All functions available in the 'Edit' menu and other settings can be accessed through this menu shortcut. It is also possible to post-edit several pages simultaneously (as long as they haven't been archived yet). Please take a closer look at the dialog box:



File size can be changed retrospectively under **'Scale/size'.** 'Scale (change resolution)' recalculates the page's physical size. 'Change size' changes the height or width of the page, which means that an empty space is created or part of the original page is truncated. 'Width' and 'height' parameters are usually entered as pixels unless the option **'Entries in percent'** is ticked. If **'Calc. proportional value'** is ticked, the entered width and height data are synchronized, i.e. changing the width would proportionately change the height as well.

'Bits & Pixel per inch': Try changing color depth or image resolution. In terms of **color depth,** '24 Bit (color)' is highest, '1 bit (b/w)' lowest. The resolution of color and grayscale images can be adjusted independently from the resolution of black/white images, however this is only applicable when editing several pages simultaneously.

'Rotation' of pages is fairly straightforward and a detailed description of the 'effects' goes beyond the purpose of this manual. Should you be unfamiliar with these **effects,** simply perform a few experiments with a test document, or consult an image editing manual. Please note that not all effects may be visible in

black/white images ('Gamma correction', for example, only affects color and grayscale images).

'Repeat effects for following number of pages' lets you apply all the parameters set in this dialog box for any number of pages, i.e. as a batch action.

➲ Note: There is no Undo function for 'Repeat effects...'!

33.5.6 Cancel last action (F5)

You can undo each of the functions in menu 'Edit' but only the function executed at the very last. The simplest way to do that is pressing the **function key 'F5'.**

33.6 Menu 'Viewing'

Tutorial on Menu 'Viewing' (movie 2min) *5.2.3*

With this menu you determine the viewing size, the quality of a page (low or high resolution) and the measures. You can also change from the viewing of one page to the viewing of two pages.

33.6.1 Normal (according to window) (F6)

This is the default setting. The page size is adapted to the screen, the page is always shown in full and depending on the size of the scanned sheet it has been either diminished or enlarged. This is the best way to get a quick overview of the pages in your archive.

33.6.2 25% of the original

The page is shown as a copy diminished to 25%.

33.6.3 40% of the original (Ctrl+F6)

The page is shown as a copy diminished to 40%.

33.6.4 50% of the original (Shift+F6)

The page is shown as diminished by half.

33.6.5 75% of the original (Ctrl+F7)

The page is shown as a copy diminished to 75%.

33.6.6 100% of the original (Shift+F7)

The page is displayed in the original size. If in the same menu the command 'As pixel equivalent (without scaling)' is activated a pixel on the screen corresponds to a pixel on the original page. If 'As pixel equivalent (without scaling)' is not ticked 1 inch on screen corresponds to 1 inch on the page. If the size of the sheet does not correspond to the size of the page on screen, go to 'Parameters' in table mode where what is displayed can be calibrated for each user.

➲ Please take note of the fact that screen copies are physically made smaller and therefore are also shown in diminished form.

33.6.7 200% of the original (Shift+Ctrl+F6)

The page displayed is the double size of the original.

33.6.8 300% of the original

The page displayed is three times as large as the original.

33.6.9 400% of the original

The page displayed is four times as large as the original.

33.6.10 Free choice in percent (Ctrl+F)

Enter the desired zoom factor in percent. The page will be shown accordingly. Possible values are from 10 to 800 percent. This setting (i.e. the zoom factor) is kept until you enter a new value or quit 'Page view'.

33.6.11 As pixel equivalent (without scaling) (Ctrl+O)

Normally the size of the page displayed is in accordance with the size of the physical sheet of paper, i.e. it is scaled according to sheet size. By running this option you determine that the scaling should take place in relation to pixels. '100% of the original...' then means that each pixel of the image file corresponds to a pixel on the screen.

➲ This option does not influence 'normal' 'Viewing' because there the page is shown according to the size of the window.

33.6.12 High resolution (printer) activated (F7)

There are two variants of each archived page. The **original copy** which possesses high resolution and the **screen copy** with a lower resolution above all designed for quick reference. The original copy is primarily needed for further use, particularly for printing. The screen copy is very useful whenever you want to go through your archive quickly. Depending on whether this menu item is ticked the high resolution copy is activated or not. You can also check the toolbar (to the right of the icons): if the filename starts with an 'A' the high resolution copy is loaded, if it starts with a 'B' the screen copy is!

➲ To print or export pages in high resolution quality you must activate the high resolution, i.e. tick this menu item.

33.6.13 Measures

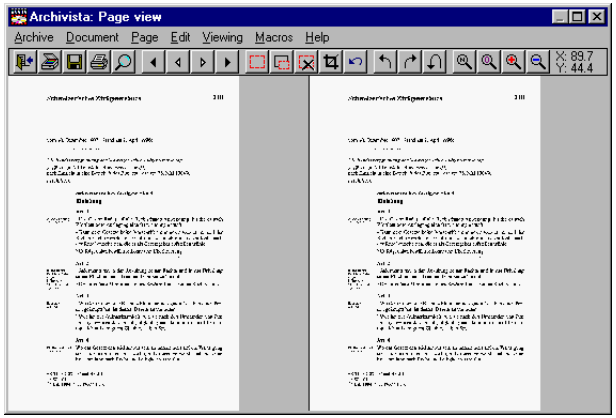
When going to 'Page view' for the first time the used measuring unit depends on what is defined in 'General program parameters' (see chapter 35.2.3). Temporarily, i.e. for the current session, you can change the measure. Possible units are: twips, points, inches, millimeters and centimeters. All entries made when in 'Page view' (including the scanning dialog box) must be made in this measure.

For your information: One inch corresponds to 1440 twips, 72 points, 25,4 millimeters and 2,54 centimeters, respectively.

33.6.14 Show two pages (F11)

Tutorial on Show two pages (movie 2min) *5.2.3*

To comfortably go through your pages you can look at two at a time. Run this menu item and you will get something like the following (in the beginning the same page appears twice):



In spite of double page mode only one page is active, i.e. you can modify only one. By default it is the right one and the pages are shown in pairs. See below for further details.

33.6.15 Left page activated (Shift+F11)

This menu item activates the left side. You achieve the same result by left-clicking the left side with your mouse.

33.6.16 Right page activated (Ctrl+F11)

This menu item activates the right side. You achieve the same result by left-clicking the right side with your mouse.

33.6.17 Show in pairs (Ctrl+Shift+F11)

When this menu item is ticked the pages are shown in pairs. Browsing takes place one by one with neighboring pages forming pairs.

➲ If this menu item is not active nor the menu item 'Show as in a book', then the inactive page remains 'frozen' where it is while the active side changes.

33.6.18 Show as in a book (Ctrl+Shift+F12)

This menu item deactivates 'Show in pairs' and enables you to browse as in a book. When you have finished looking at pages 2 and 3 and you click the icon 'Next page in document' pages 4 and 5 are displayed.

33.6.19 Additional icons

On running this menu item a number of additional icons will be displayed, which simplify manipulation of pages (deleting, moving, importing, exporting and post-editing them).

33.7 Menu 'Macros'

Running macros is possible also within 'Page view'. They work the same as in table mode. Please go to chapter 35.8 to see how macros are edited.

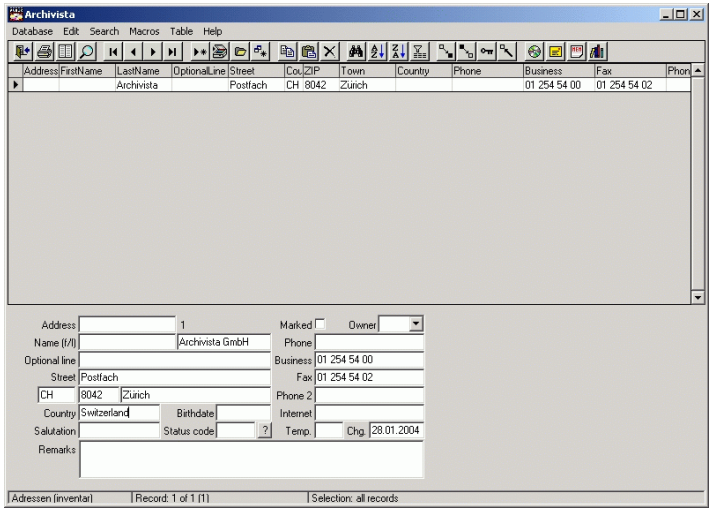
34 Further Tables

Beside table 'Archive' there are other tables in Archivista 2009/I, which are accessible via menu 'Table'. Please note that working with these tables is similar to working with table 'Archive', however, a few menu items are missing. What is more, the items in menu 'Edit' often carry the name 'record' instead of 'document'.

➲ You can work with the following tables only when they are opened during the start up process. For this purpose the option 'All tables' under 'Parameters'/'General' must be activated.

34.1 Table 'Addresses'

Table 'Addresses' offers you comfortable address management.


Archivista
Database Edit Search Macros Table Help
Address FirstName LastName OptionalLine Street Cou ZIP Town Country Phone Business Fax
Archivista Postfach CH 8042 Zürich 01 254 54 00 01 254 54 02
Address 1 Marked Owner
Name (f/l) Archivista GmbH Phone
Optional line Business 01 254 54 00
Street Postfach Fax 01 254 54 02
CH 8042 Zürich Phone 2
Country Switzerland Birthdate Internet
Salutation Status code ? Temp. Chg. 28.01.2004
Remarks
Adressen (inventar) Record: 1 of 1 (1) Selection: all records

34.1.1 Fields 'Country code, 'Zip code', 'Town' and 'Country'

You can make good use of the editing function 'Get abbreviation' in the address fields 'Zip code' and 'Town'. When you run it in these fields the program automatically fills in the remaining fields 'Country code', 'Zip code', 'Town' and the name of the country. The program gets this information from the auxiliary table 'Zip codes (addresses)'.

In the following an example: You are in the field 'Town' and enter 'a'+'F2'. With the Swiss zip codes tied up, the program immediately recognizes this as 'Aarau' because it is the first town beginning with 'a' in the zip code list. Consequently the field 'Town' is completed to 'Aarau' and the field 'Zip code' to '5000'. If you work with more than one list of country zip codes, it will make sense to enter the country code too (plus a hyphen): press 'ch-a' + 'F2' and also the fields 'country code' and 'country' will be filled automatically. Conversely, when you enter the zip code and press 'F2', Archivista will enter the name of the town.

The other fields in table 'Addresses' behave 'normally', i.e. the function 'Get abbreviation' ('F2') refers to the auxiliary table 'Abbreviations' (and not to the table 'Zip codes `[addresses]`').

➲ Please make sure that the auxiliary table 'Zip codes (Addresses)' contains at least one post or zip code directory. Only if there is a directory you can profit from Archivista's clever post code function.

34.1.2 Status code

On the right of the field 'Status code' there is a question mark. When you click this you get a message regarding the conventions of the status codes. You can adapt these yourself under 'Parameters' – under the condition that you enjoy systems operator privileges.

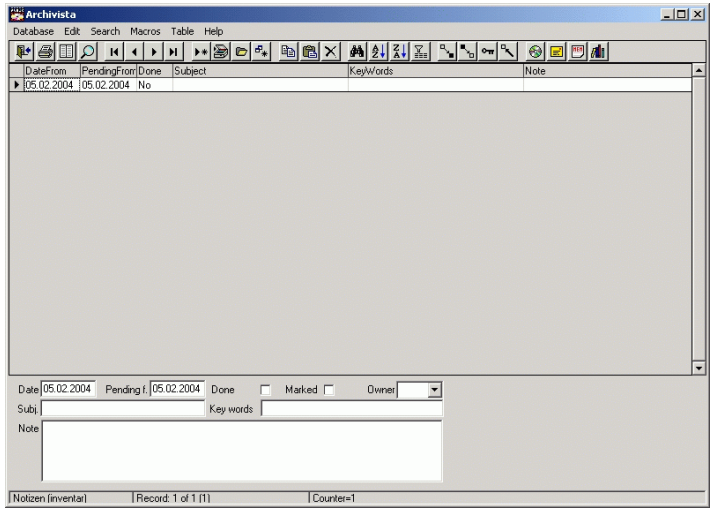
This field should be seen as a sort of multi allocation field with the help of letters. In the following an example. Let us assume that you have defined the following status codes: W=work, F=family

and P=picture post cards. When you create a new address you must decide whether the person is one of your work mates, belongs to your family and whether s/he is to get a picture postcard from your holidays. For each 'yes' you enter the relevant code in the status code field of your newly created address, i.e. your aunt who is to get picture post cards from you is marked with an 'F' and a 'P' in the status code field. (The order of these letters does not play a role.)

Later you can select your addresses according to a status code. Before your holidays, for example, you can search for all the addresses that you need to take with you to write post cards. (Go to 'Selection', 'Status'. Under variable enter 'P'.)

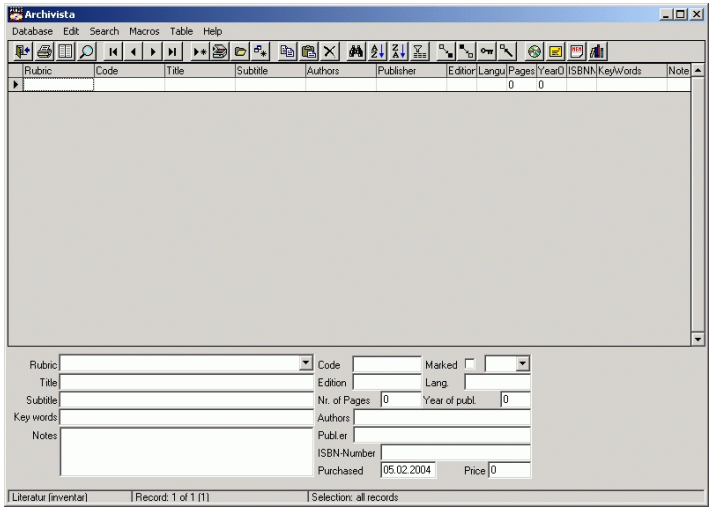
34.2 Table 'Notes'

Table 'Notes' serves to manage notes. The layout of this table is extremely simple and needs no further explaining.


Archivista
Database Edit Search Macros Table Help
DateFrom PendingFrom Done Subject KeyWords Note
05.02.2004 05.02.2004 No
Date 05.02.2004 Pending f. 05.02.2004 Done Marked Owner
Subj. Key words
Note
Notizen (inventar) Record: 1 of 1 (1) Counter=1

34.3 Table 'Literature'

Table 'Literature' helps you to keep an overview of books and other publications.


Archivista
Database Edit Search Macros Table Help
Rubric Code Title Subtitle Authors Publisher Editior Langu Pages YearO ISBN KeyWords Note
Rubric Code Marked
Title Edition Lang.
Subtitle Nr. of Pages Year of publ.
Key words Authors
Notes Publ.er
ISBN-Number
Purchased 05.02.2004 Price
Literatur (inventar) Record: 1 of 1 (1) Selection: all records

The fields 'Rubric' and 'Code' (=call number) as well as the auxiliary table 'Rubric (Literature)' are particular to this table.

34.3.1 Fields 'Rubric' and 'Code'

It makes sense to allocate your books to certain rubrics such as 'economics', 'law' or 'Elizabethan literature.' To keep order in your rubrics and to know at all times with which rubrics you work you can access a list of all existing rubrics in the field 'Rubric'.

To allocate a rubric to a record choose the desired entry in field 'Rubric' and quit the field. In field 'Code' the corresponding code or call number appears automatically. For a fully-fledged call number system you now must edit the field code and add a few numbers to make your call number unequivocal.

34.3.2 Auxiliary table 'Rubric (Literature)'

The rubrics for table 'literature' are defined in the auxiliary table 'Rubrics (literature)'. Go to menu 'Table' and the corresponding menu item. There you add new entries, modify and / or delete them. The entry you get later in table 'Literature' is taken from this auxiliary table, i.e. the final entry depends directly on what you define here.

Part IX

RichAdmin

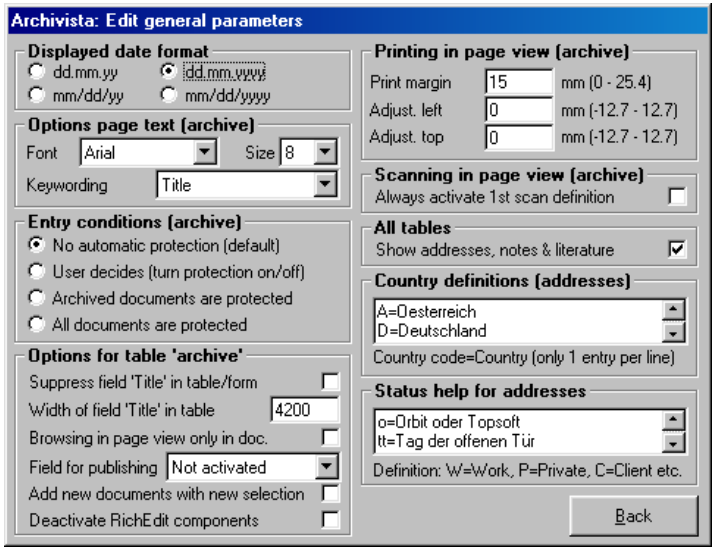
35 Parameters

Please be aware of the fact that when you venture into making parameter changes you might massively change the program's basic workings. You need 'SYSOP' privileges to change parameters.

On a first level we differentiate between general parameters, user administration, parameters for archiving documents, selection parameters and macros. Defining print reports contains complex functions which is the reason why we have made a separate chapter (see chapter 37)

➲ Parameters in Archivista are set per database and also stored with the database. Whatever you change in 'Parameters', the changes take effect only in the currently active database.

35.1 General program parameters



35.1.1 Displayed date format

The default setting is 'mm/dd/yy' which means that date fields show a two digit format with the month appearing first. The 13th of March 1999 will read '03/13/99', for example. You can also go for a four digit year ('mm/dd/yyyy') or, if you prefer a European format, use either of the two remaining formats ('dd.mm.yy' or 'dd.mm.yyyy').

35.1.2 Options page text (Archive)

Under **'Font'** and **'Size'** you specify font and size for the field 'Page text' in table mode of table 'Archive'. Thanks to this option 8 bit fonts of foreign languages may be installed.

Under **'Keywording'** you determine the target field of information from 'Page text' that you single out as keyword information worthy to be transferred to a different field. As an example may serve the following: in 'Page text' you select the word 'competition' to be copied and by pressing function key 'F2' you paste it automatically in the field specified here under 'Keywording' (e.g. field 'title').

35.1.3 Entry conditions table 'Archive'

When you alter the content of a field in the main view of Archivista the changed content will normally be saved without further ado. This corresponds to the first entry condition under this heading which is the default specification.

If you turn one of the other options on, it will have an impact on the main view of Archivista: close to the check button 'Activated' at the bottom of the window a new check button 'Protected' will appear. Depending on the situation you can enforce with

this option that a user must confirm his or her entry before it is definitively saved.

➲ Please note that users with permission to alter documents will still be able to do so but if the check button 'Protected' is turned on they will have to confirm changes. In the following the different options are described:

User decides (turn protection on/off): the check button 'Protected' appears; it is the users' decision to turn it on in order to avoid inadvertent changes of field information; all changes require confirmation.

Archived documents are protected: the check button 'Protected' is turned on when you are attempting to edit a document that has already been archived.

All documents are protected: Even if you untick 'Protected' while editing a document, the option will be turned on again when you leave the document in question. Only for the freshly created document the option is turned off.

➲ You can turn the button 'Protected' on and off by using **function key 'F12'**.

35.1.4 Options for table 'archive'

With the first option **Suppress field 'Title' in table/form** you decide globally (i.e. for all users) whether the field 'title' should be displayed or not. This change takes effect as soon as the next Archivista client is started up.

Width of field 'Title' in table allows you to determine the width of the field 'Title' in the table of the main view.

By checking the third option **Browsing in page view only in doc.** you restrict the browsing to the currently active document, i.e. when using the functions 'next page' and 'previous page' you can no longer move beyond the borders of the document.

The field that you specify to the right of **Field for publishing** is a kind of second owner field with which a document can be allocated to a second user or user group. For many archives this is an irrelevant function. However, it can turn out to be important in some cases. Take the following instance:

A knowledge archive is continually fed with new documents by two dozen project managers. Each project manager is responsible for the documents she added. At the same time she can make the documents available to an extended circle of managers.

After this option has been activated the input mask will show a second combo box with document owners. This way each primary owner of a document can set a second owner. It goes without saying that the main user settings apply also in the extended owner environment. See chapter 35.2.

➲ The field that you specify here must be of the type 'Definition'. See 36.6.3.

The last option in this frame, **Deactivate RichEdit components,** serves to turn off formatting (e.g. italics) for memo fields. This function was developed for companies with subsidiaries in different countries. Depending on which language the keyboard works with character font incompatibilities might appear. This option circumvents the problem.

35.1.5 Printing in 'Page view' (Archive)

Within 'Page view' (only table 'Archive') there is the possibility to send pages to the printer. Since some printing devices have margin settings that unduly diminish the printable space of a sheet it may turn out useful to make up for them here. What is more, you can adjust the starting position on the left and at the top.

35.1.6 All tables

With this option you specify whether the tables 'Addresses', 'Notes' and 'Literature' (as well as their auxiliary tables) are opened when Archivista is started up. If this option is not turned on, starting the program is faster but you cannot access the mentioned tables.

35.1.7 Country definitions

Country definitions are exclusively needed for table 'Addresses' and table 'Zip codes (Addresses)'. This field contains the country definition for the field 'country' in table 'Addresses' and is the link to the country codes in table 'Zip codes (Addresses)'. An example: if you want the field 'country' to say 'Suiza' instead of 'Switzerland', it suffices if you enter 'CH=Suiza' here.

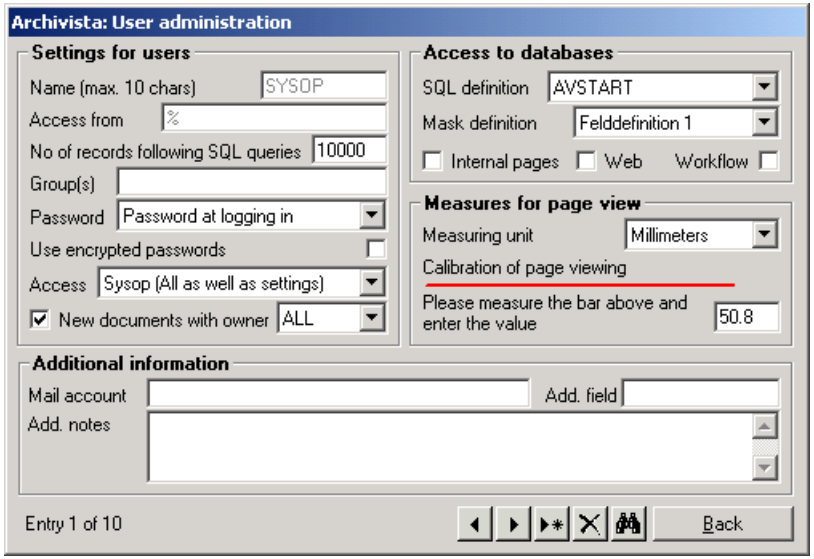
35.1.8 Status help for addresses

This is the location to specify the reference list for your status letters (table 'Addresses'). More information you find in chapter 34.1.2.

35.2 User administration

Tutorial on User Administration (movie 2min) *5.5.1*

All Archivista 2009/I settings referring to users you find in this form.



➲ There is a difference between the number of licenses and the number of defined users. Even with only one license of Archivista 2009/I a virtually unlimited number of users can be defined. The individual users cannot work with the system at the same time, though. Only as many persons can work at a time as licenses have been bought.

When an Archivista database is created solely the user **'SYSOP'** is created along with it. This is the default user and equipped with sysop privileges, i.e. she can run all existing functions in Archivista. No password is set yet and no logon dialog box appears.

Only after the entries mentioned below have been made the system will ask for a password any time the program is opened.

To create users, to select or delete them there are five buttons at the bottom right of the dialog box. (To the left of these buttons you see which user is active at the moment).

➲ **Create new user:** Please note that when creating a user you always have to enter 'Name' and 'Access from' (e.g.

'uer@localhost'). More information on this can be found under 35.2.1.1 und 35.2.1.2.
The following entries must be made for each user:

35.2.1 Settings for users

35.2.1.1 Name

Each user needs an identification name, i.e. a non-equivocal term which establishes a person as a rightful user and which gives him individual privileges to use the database. A name may consist of a maximum of 10 letters and numbers but it cannot begin with a number ('4DOS' is not a valid name but 'DOS4' is).
➲ User name and owner are identical insofar as each user is listed as owner in the field of the same name (see chapter 29.7.2). However, a user can allocate a different owner to a record (e.g. a document) under the condition that he has the necessary privileges.

35.2.1.2 Access from

Please specify here from where a user may access the Archivista database. Under 'Access from' you enter either 'localhost', '%', a domain like company.com (also possible in combination with '%') or an IP address such as 192.168.0.44 (also possible in combination with '%').
➲ Please note that you have to use 'localhost' for all the users who work on the PC on which MySQL is installed. If you want to access the Archivista database from a different PC, then you need either global access rights with '%', the domain name of the computer accessing the database or the IP address of the computer.

35.2.1.3 Number of records following SQL queries

The advantage of client/server solutions lies in the fact that database queries (such as "we want to see all documents of a particular customer from last year") are not dealt with locally by the end-user's PC, i.e. the "client", but by the database server. The server processes the search and the client receives only the results.
Theoretically, a client could ask for all the records. In other words, for the entire database. The transferral of this mass of information to the client PC might take quite long in this case (especially with very large databases). This is the reason why we specify the maximum of records transferred from server to client for each query.
This value is entered under 'Number of records following SQL queries'. It corresponds with the number of list items to be seen by the user in the main view of the program.
➲ When a query is sent to the server the following happens: in addition to supplying the list items the server names the total number of records resulting from the search query in order that the enduser may know whether the results she sees constitute a subset or a total of the results she produced by her query.

35.2.1.4 Group(s)

In principle, 'Groups' in Archivista are ordinary users. The only difference is that 'Groups' are allocated to users.
To each user you can allocate one or more **'Groups'**. The user in question may see and edit not only his own documents but also those of the group(s) as if they were his own.
Example: Let us assume that users A, B and C form a group. D is not to have access to the documents of A, B, and C. S possesses all privileges but his documents must not be viewed by either A, B, C or D. With the function 'Group(s)' this riddling task is elegantly solved. You create a virtual user ABC which is allocated to the users A, B and C (but not to D!). In this way these users form a group: A, B and C each view their own individual documents

as well as those of user ABC, i.e. the 'Group.' D does not have any access and S has access to all documents anyway since he has reading and writing privileges with regard to all documents.

➲ From version 5.1 onward each user may belong to several groups. You achieve this by entering the names of the groups separated by commas. For example, if you want a user to have access to the documents of both 'Sales' and 'Purchasing' you simply enter 'Sales,Purchasing' in the field 'Group(s).'

The field 'Groups' must also be used in combination with the option 'Field for publishing'. See chapter 35.1.4 and 35.2.5. When users must be placed in a position to allocate subordinate owners (or owner groups) to their documents, then, the user (or group) to which ownership must be extendable must be named under 'Groups' in the dialog box 'User administration.

35.2.1.5 Password

There are the following three possibilities:

Without password: the user possesses no password and is allowed to login without identifying himself beyond giving his name.

Password at login: During login a password must be entered. Without a password no login is possible.

New password at login: On the occasion of the next login the system asks first for the old password (if one was defined). After that, Archivista 2009/I asks for a new password which must be confirmed. By choosing this possibility the systems operator can force her users to set a new password. The new password must be different from the old one.

➲ In contrast to the user name the password may consist of numbers only (a valid password might be '56777', for example). Password administration takes place directly in MySQL, i.e. you can change passwords directly within the MySQL console program (or a third-party program). Please note further that a password change takes effect in all databases since MySQL manages passwords globally.

35.2.1.6 Using encrypted passwords

On the basis of the rights that are allocated to users in Archivista (see 35.2.1.7 and 35.2.1.4) a user is usually not allowed to see each and every document. This access authorization feature is 100% secure with Archivista clients but not on the level of the MySQL database. In MySQL access authorization rules cannot be easily laid down on the level of individual records because version 4.0.x of MySQL does not know any Views yet and Merged-Tables exist only in Beta quality.

As a result, users might use their Archivista password to log into an Archivista database via a third party program (e.g. mysql client program) and look at all the records. In an environment where this is undesirable, i.e. where access authorization is taken seriously, Archivista administrators should tick 'Use encrypted passwords'. This way log-in is only possible via an Archivista client but not via a third party application.

➲ This option cannot be activated for the first user (SYSOP). However, it would not make much sense because the Sysop has all rights anyway.

35.2.1.7 Access

Under 'Access rights' you define which records within the database the user is allowed to work with. We differentiate among the following:

Read public ones, edit none: Only those records are shown which do not have a specific owner (or who belong to the owner 'ALL'). No data may be edited.

Read public ones, edit own records: The records not belonging to any specific user and the records belonging to the user we are talking about are shown. The latter may be edited.

Read all, edit own records: All records are shown, i.e. also those belonging to other users. But those belonging to the other users cannot be edited.

Read all, edit all: All records are shown and may be edited.

Sysop privileges (all as well as settings): All records are shown and can be edited. In addition to that the user with sysop privileges has access to all settings and parameters.

➲ Imagine the SYSOP as a sort of super user because she is mistress over the rights of all other users. Normally you login as SYSOP only when you want to change parameters.

35.2.1.8 New documents with owner

Tutorial on New Documents with Owner (movie 2min) *5.5.2*

Here you can specify whether a user may create new documents and if yes, which user they should be allocated to.

➲ If a user has no right to create new documents she has no right to delete documents either.

35.2.2 Access to databases

Under **'SQL-Definition'** you can specify which document selection a user should see after starting up Archivista. The definition itself is made in the 'Definitions for selection,' see chapter 35.7.

The option **'Mask definition'** serves to specify the input mask a user sees after starting up Archivista. The definitions for the input masks are made where the fields themselves are created. By being able to select different input masks for different users you can have several archives in one.

When the check-box **'Internal pages'** is ticked the bitmap pages are delivered straight from the database. When you leave it unticked (default) the bitmaps are loaded from the directories. The advantage of this (ticked) option is that users no longer need read access to the directories of your archive: it helps you to simplify user administration and increase security.

➲ Please note that users for which this option is ticked cannot create new documents. This option makes most sense for users who view and edit documents by means of the web browser and who do not create documents manually.

The check-box **'Web'** enables user access via WebClient. If it is unticked the user cannot access the archive database via the Web module. See also 9.

By means of the check-box **'Workflow'** you can trigger certain actions when certain events (creation of new document, alteration of document, etc.) take place. Please refer to 8.6.2 for more information.

35.2.3 Measures for page view

Possible measures are: twips, points, inches, millimeters and centimeters. You can define for each user with which measuring unit they may work.

For your information: one inch corresponds to 1440 twips, 72 points, 25,4 millimeters or 2,54 centimeters, respectively.

'Calibration of page viewing' is needed to adapt the display of pages for each user or for each PC screen. Do make the calibration directly on screen! Measure the red bar with a ruler and enter the resulting value according to the chosen measuring unit. From now on the size of the pages is displayed in accordance with the calibration, i.e. the size 'US letter' is 'US letter' also on screen.

35.2.4 Additional information

You can collect one e-mail address per user. The e-mail address may be used by third-party modules (e.g. workflow module) when

certain events trigger user notification. 'Add. field' means that you can allocate an additional identification number to a user, if need be, and in the memo field 'Add. notes' you can enter any text you like.

35.2.5 Working with owners

35.2.5.1 Working with one owner

The user administration of Archivista may sometimes seem confusing because a series of options can be tied to the user name and they can be combined in different ways.

This subchapter serves to clarify the user and owner concept of Archivista. First we discuss how the options 'Access', 'New documents with owner' and 'Group(s)' interact. Then, the following subchapter addresses the possibility of allocating a second document owner.

Uppermost in the hierarchy of what a user may or may not do with regard to a document is the option **Access** as described in 35.2.1.7. The option **New documents with owner** (35.2.1.8) may seem uninspiring but with its help you can allocate a document to a new user straight after creation and make the creator lose ownership. The option **Group(s)** means nothing else than if there is an entry, then the user's rights are extended to someone else's documents.

The following table may illustrate how these options are best combined. Because of space reasons options were abbreviated. The column 'Group(s)' does not give any suggestions for entries but it simply states whether an entry is compulsory or not.

	Entry under 'New docs w. o.'	**Entry under Group(s)**
R. public., edit none	`[cannot be combined]`	not compulsory
R. public., edit own	own user	not compulsory
	ALL	not compulsory
	other user	compulsory
Read all, edit own	own user	not compulsory
	ALL	not compulsory
	other user	compulsory
Read all, edit all	own user	not compulsory
	ALL	not compulsory
	other user	compulsory

It goes without saying that not all of the above combinations are equally sensible. If a user can read and edit all documents, why should she be part of a group then? Nevertheless, the overview helps to make sure one has the right combination.

As a typical example of the allocation of detailed user rights may serve the following company.

The accounting department of a medium sized company employs

- Ruth and Rudy in Accounts Receivable (ruth and rudy)
- Peter and Pauline in Accounts Payable (peter and pauline)
- one chief accountant (c)
- one apprentice in Accounts Receivable (ARa)

What is a good way to allocate rights?

Solution:

User		New doc with o.	Group(s)
ruth	R. public, edit own	AR	AR
rudy	R. public, edit own	AR	AR
peter	R. public, edit own	AP	AP
pauline	R. public, edit own	AP	AP
c	Read all, edit all	c	[none]
ARa	R. public, edit none	[none]	AR

When Ruth creates a new document it immediately gets the owner 'AR' and since both Ruth and Rudy belong to the group AR, each can treat the document as if it was their own. The chief accountant may view and edit all documents. The apprentice in Accounts Receivable may only **view** documents, namely the public ones and those belonging to the department Accounts Receivable.

The basic user settings do not change if a user belongs to one or more groups. That is, Ruth who has the rights 'Read public ones, edit own' cannot only view the documents of the group 'AR' but can also edit them.

➲ Please take care to set up group users **before** allocating a user to a group.

35.2.5.2 Working with two owners

As long as MySQL does not possess 'Views' functionality one must make use of a trick to be able to allocate **two** owners to a document. Although completely unnecessary in many cases this option can be very valuable, for example in the following framework.

Project managers add documents to a knowledge archive on a daily basis. Each project manager remains responsible for those documents added by her. At the same time she can make the documents available to an extended circle of users. To activate the second owner there are three steps to go through.

- First and second step: create field and specify it as second owner field; see 35.1.4
- Third step: Determine in user administration under 'Group(s)' who may assign which subordinate user(s); see 35.2.1.4

Now users have a field with which to 'publish' their document(s) to other users.

In our example user 'freuler' can allocate the subordinate user 'ALL' to her document. Alternatively, she can publish it to the sales team only (group SALES).

Needless to say, the basic user settings apply also in the environment of extended ownership. A document that has been assigned to me as second owner I can look upon as 'my own document'. Users that are allowed to edit their 'own' documents can also edit documents of which they have subordinate ownership only.

It is even possible to have more rights as second owner than as first owner. Let us take a look at the following example: A creates a document and straight after this selects ABC from the owner pull-down field (ABC possesses the rights 'Read public ones, edit none'). Then, A assigns the document to B by means of the second owner field. Since B possesses the rights 'Read public ones, edit own' B will be able to make changes to the document while ABC

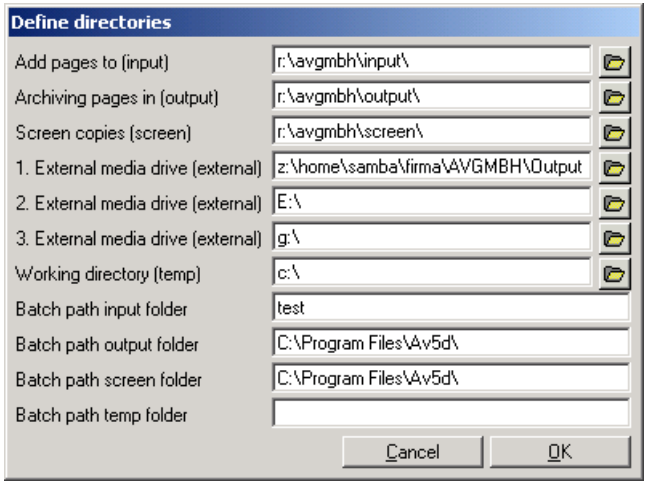
will not. Can A still edit the document later? Yes! Although A is no longer owner of the document, he is part of group ABC and since he has editing rights he cannot only view the document but also change it.

35.3 Directories (folders)

Every Archivista database possesses a number of directories, the administration of which is done with this menu item.

35.3.1 Archivista directories

Only the status information of the archive is stored in the database itself. The reason is that the database does not become huge and unflexible. The pages belonging to the documents and the index files are stored in separate directories (exception: internal pages). We differentiate between the following directories (in brackets the path names for the first database if Archivista has been installed in 'C:\Program Files\Av5e'):



Added pages (C:\Program Files\Av5e\archive\input): All pages added via 'Page view' are filed in this directory if they do not go straight into the database. Since the folder number is not definitive at this point the final document folder does not exist yet either. The directory named INPUT is the relevant directory for added pages before they are archived. For the sake of optimized filing the Input directory contains 100 subfolders (in000 to in099). The documents are allocated to these subfolders on the basis of their modulo value with regard to 100.

Archived pages (C:\Program Files\Av5e\archive\output): When they are archived they are filed in this directory. Now that the folder number has been determined filing happens in the corresponding document folder (e.g. document 455 in folder '3' gets filed in 'c:\Program Files\Av5e\output\archive\arch0003').

Screen copies (C:\Program Files\Av5e\archive\screen): Screen copies are optional. When chosen filing happens in analogy to the filing of archived pages. Filing of screen copies takes place into a directory named 'screen'. Only the screen copy gets filed in the corresponding folder (e.g. 'C:\Program Files\Av5e\screen\archive\arch0003').

Index ('C:\Program Files\Av5e\archive\index'): This directory houses the index of the archived documents and pages. Remember: the index is created only if we treat our pages with OCR and work with ASCII files to have a full text retrieval archive.

For external media ('C:\', 'D:\' and 'E:'): In correspondence with these settings pages already transferred to other media can be addressed. In most cases these other media will be CDRs which you insert in one of the mentioned drives.

35.3.2 Temporary working directory

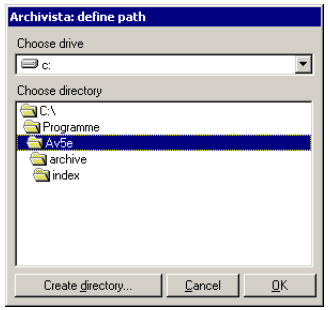
Specify a directory which you want to use for temporary data transfer. Choose a directory of the fastest drive in order to achieve maximum velocity of Archivista 2009/I.

35.3.3 Batch paths

For the Business and Enterprise versions directory paths can be specified which allow for the scanning via a network document server.

35.3.4 Archivista: Define directories

The standard procedure for entering a directory including its path is to click the button to the right of the field. A form will appear which serves to comfortably choose the directory in which your files (the image files, the screen copies, the ASCII files (=the text recognized by the OCR software), etc.) should be put. After choosing the directory you confirm your entry by clicking 'OK'.

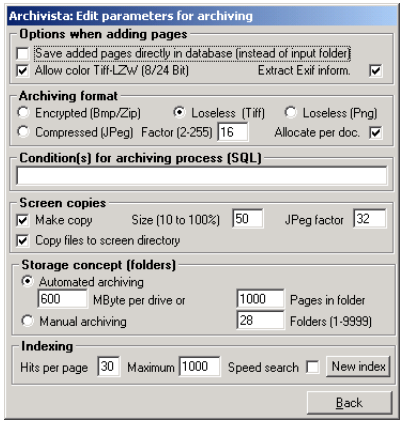


➲ If you want to assign UNC paths, you need to do it manually. UNC paths cannot be assigned with the help of the little browse buttons.

➲ Please note that a directory may be used only by one Archivista database. If you allocate the same directory to more than one Archivista database, you will get an error message.

35.4 Archiving & index...

In the submenu 'Archiving & index...' you find parameters, which are important to archiving and transferring copies of pages to external media.



35.4.1 Options when adding pages

The option **Save added pages directly in database (instead of input folder)** saves you the security risk of having to give users access to server directories.

By ticking the option **Allow color Tiff-LZW (8/24 Bit)** also color pages can be stored in Tiff format. Since the year 2004 Tiff-LZW is freely available, i.e. previous patents expired at that point in time and there is one more alternative for color pages.

The option **Extract Exif information** serves to extract Exif information when digital camera pictures are imported to Archivista 2009/I.

35.4.2 Archiving Format

Under this heading you specify the format in which pages are saved. The following four formats are currently available:

Encrypted (Bmp/Zip): This format consists of Windows bitmap files which get compressed in zip format. The fully automated encryption of your data is a unique feature. Each single page (black and white, gray scales or color) is encrypted and put on

your hard disk. It can only be accessed via Archivista 2009/I – under the condition that you have the necessary privileges.

Loseless (Tiff): A common filing format in the world of document management. In black and white pages need little disk space (at 300dpi ca. 50 KByte). By ticking the option 'Allow color Tiff-LZW (8/24bit)' (see previous section) the Tiff format can also be used for color pages.

Loseless (Png): Archivista's default format, good compression as regards black and white pages, pages in gray scales and in color. This format is freely useable and is supported by Internet browsers.

Compressed (JPeg): Excellent compression rate but not loseless, i.e. the compressed pages are not of the same good quality as in the original (in contradistinction to the three formats above. The **JPeg format** makes most sense where there are **color pictures** to be archived.

You must specify a compression factor (2 to 255) for the format 'Compressed (JPeg)'. Please note that this factor may be entered separately for the original, high resolution page (in 'Archiving format') and the screen copy (in 'Screen copies'/JPeg factor)

➲ Different JPeg values for compression under 'Archiving format' and 'Screen copies' mean that the original copy and the low resolution screen copy are archived in different quality. Since the screen copy is by definition of lower quality its JPeg-factor should be higher than the one under 'Archiving format'! A high JPeg-factor means lower quality but an attractively small file size.

35.4.2.1 Allocate per document

Archives which were created with version 4.14 (or lower) allowed you to choose the archiving format only on the level of the database and before adding documents and pages. If the option 'Allocate per doc.' is not activated, Archivista 2009/I will react in the same way as earlier versions, i.e. you will not be able to change the archiving format on document level. The option must be ticked in order that field 'Format' may be changed upon creation of a new document.

35.4.3 Conditions for archiving process

Here you can specify by means of an SQL fragment that the archiving process should only involve certain documents and not others.

35.4.4 Screen copies

By ticking **'Make copy'** you determine that a screen copy should be created. This option is ticked by default and you should keep it like that because the pages you are going to add are displayed much faster later.

Under **'Size (10-100%)'** you can enter the size that the screen copy will take in relation to the original, high resolution copy. Normally you will work best with a value in between 25 and 50 percent (50% is the default value) – it depends on the material you want to scan.

The option 'JPeg factor' enables you to determine the quality of the screen copies for documents which exist as jpeg files. A high value means small file size but also lower quality.

The field **'Copy files to screen directory'** serves to determine whether the screen copies should be copied to the directory or not.

➲ Please be aware of the difference between 'Make copy' and 'Copy files to screen directory'. The first option creates a screen copy in the 'output' directory. The second causes an additional copy to be stored in the 'screen' directory.

35.4.5 Storage concept (folders)

Normally, you work with **'automated archiving',** i.e. Archivista creates and handles the necessary directories. As long as you work

with CDRs (write-once CDs) for external storage media you should not change any settings.

If, however, you work with a different external storage medium, you must specify the number of **'MBytes per drive'.** In addition you can specify the number of **'Pages in folder'.** This is the number of pages which should go into a directory when the pages are archived. With regard to the pages please note that two files are created per page (original page and screen copy).

➲ The field 'Folders' (= folder number) is an incremental field, i.e. the number appearing there is automatically increased when enough pages have been added and archived.

If you opt for **manual archiving,** you determine yourself when a new folder should be created. You do that by entering a folder number increased by one in the field **'Folders'**. All subsequent files are filed in a new folder (or put in a new folder upon creation). Once a folder number is increased within the option 'Manual archiving' it is increased for good. You cannot change the folder number back. We recommend that you increase the folder number only after due consideration. The option 'Manual archiving' makes sense only in exceptional cases.

35.4.5.1 Indexing with MySQL

When this option is turned on indexing happens continually in the background in a fully automated way. The index is at all times up to date.

➲ Working with MySQL indexing should not present any problems. However, please keep in mind that the creation of a totally new fulltext index may require considerable amounts of temporary space on the hard disk. How much disk space exactly is necessary depends on the pages in your archive. By rule of thumb it makes sense to reserve 1 GByte temporary disk space per 100 MByte of page texts.

35.5 Internal image copies

With the options in this dialog box you specify how the internal image copies are created. By internal image copies we mean those images which are stored directly in the database.

By ticking **'Activate internal image copies'** you effect that internal copies are made from the original high resolution files.

➲ Please note that pages in the formats BMP/Zip and TiffG4 are automatically converted to Jpeg. BMP/Zip and TiffG4 cannot normally be displayed in a Web browser.

The **Size in % of original (10-100%)** expresses the size of the internal copy in percent of the original copy. `50` means that the file size of the image copy is half the file size of the original copy.

Under **JPEG factor of image copy (2-255)** you define how much the pages are to be compressed for browser display. If the original copy is a JPEG image you need not enter anything. But if the original copy is a Tiff image it is worth adjusting the compression factor.

35.6 Field names

Due to the fact that the original application was written in German the database coded field names are in German. To ensure compatibility between the different language versions the field names are kept in German. For modifications in the 'Definitions for selection' and the 'Print reports' you have to work with the original field names. In the following you find a conversion table:

German=English

Titel=Title

Datum=Date

Akte=Document
Seiten=Pages
Stichwörter=KeyWords
Personen=Persons
Notiz=Note
ErfasstDatum=EntryDate
Ordner=Folder
Farbe=color
EDVName=FileName
Erfasst=Added
Archiviert=Archived
Eigentümer=Owner
Markiert=Marked
Laufnummer=Counter
NotizRTF=NoteRTF
Verbindungen=Hyperlinks
Gesperrt=Locked
Anrede=Address
Vorname=FirstName
Nachname=LastName
Zusatzzeile=OptionalLine
Strasse=Street
Landcode=Countrycode
PLZ=ZIP
Ort=Town
Land=Country
Telefon=Phone
Geschäft=Business
Telefax=Fax
Zusatz=Phone2
Geburtsdatum=BirthDate
Temporär=Temporary
Briefanrede=Salutation
Bemerkungen=Remarks
Aufnahme=EntryDate
BemerkungenRTF=RemarksRTF
DatumVon=DateFrom
PendentAb=PendingFrom
Erledigt=Done
Betrifft=Subject
NotizRTF=NoteRTF
Untertitel=Subtitle
Autoren=Authors
Verlag=Publisher
Auflage=Edition
Sprache=Language
AnzahlSeiten=PagesOfBook
Ausgabejahr=YearOfPublication
ISBNNummer=ISBNNumber
DatumKauf=Purchased
Preis=Price
Gebiet=Region

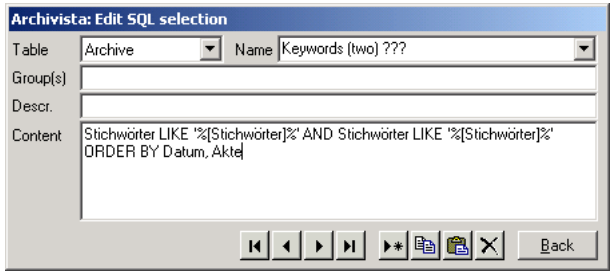
Note: The end user does not see any German field names, everything will be translated to English field names by the program.

35.7 Definitions for selection

Archivista uses an internal selection process that is based on SQL. This concerns routines behind the commands in menu 'Selection'. SQL is an abbreviation of 'Structured Query Language'.

➲ We recommend that you work through this chapter only if you have some basic knowledge of SQL or if you are ready to invest some time to make yourself familiar with SQL. If SQL is all Greek to you – as the phrase goes – you should not make any alterations in the selection criteria / filters.

Archivista offers you the possibility to define your own selections on the condition that you understand SQL. However, the formulation of selection criteria in Archivista requires only a limited number of SQL commands. Go to 'Parameters', 'Definitions for selection...' and you will get the following mask:



The table for which you want to create a new definition must appear in field 'Table'. Enter the title (i.e. the entry to appear afterwards in menu 'Selection') in field 'Name'. Field 'Content' houses the SQL definition itself. Field 'Descr.' (for 'Description) may be used to add further comments on the definition.

➲ In the field 'Group(s)' you can enter users (group users). By this you specify which users may work with which SQL definition. Users with SYSOP privileges see all definitions.

35.7.1 SQL commands used by Archivista

In the following a short description of how selection criteria or filters must be formulated in principle. The individual fragments must be separated by a space each and are to be used in the sequential order described below.

```
SELECT * FROM Table WHERE (Conditions)
ORDER BY (Sort)
```

➲ The underlined part of the SQL fragment must not be entered. It will be added automatically when a selection is made via menu 'Selection'.

(Conditions): The conditions serve to restrict the selection of records. A more detailed description would exceed the scope of this handbook. If you do not understand the following examples, please use a handbook on 'Access'. Leaving away the condition 'WHERE' would result in all records being selected. You can use the asterisk ('*') as wildcard symbol, however, you must use it as in Access SQL.

(Sort): After the fragment 'ORDER BY' the field names follow according to which the sorting takes place. You can use several fields. The abbreviations 'ASC' and 'DESC' result in ascending and descending sorting, respectively. 'ASC' is optional, i.e. when nothing is mentioned an ascending sorting is effected. If you omit the condition 'ORDER BY', no sorting will take place.

35.7.2 Extensions of SQL queries

Archivista offers two extensions to SQL queries. 1. Archivista stores all SQL queries in table 'Parameters'. 2. Variables: if you want to use a variable in a standard query ('definition for selection'), then you must put the name of the field in question in square brackets (e.g. **'[Persons]'**. When you choose this particular definition for selection, Archivista will ask you for a variable for the field **'Persons'**.

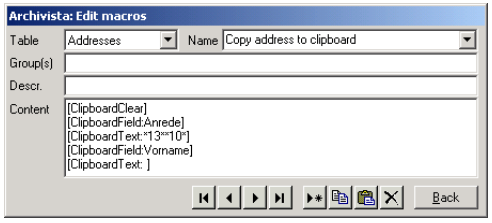
➲ You can also use curly brackets and enter {Persons }by which you effect that a variable once entered is re-used in other places of the SQL search string (e.g. when you want to search for the same term in all fields).

35.7.3 SQL selection at start of program

You may specify the kind of document selection that shall be displayed in table 'Archive' after the program has been started. Make your definition in 'AVSTART'. The illustration below shows the entry for a descending order of documents.

35.8 Edit macros...

Archivista offers you a simple macro programming language in menu 'Edit macros'. For each main table you can define 9 macros and for 'Page view' belonging to table 'Archive' there are 5 slots for macros. Go to 'Parameters', 'Edit macros'. Choose the table for which the macro should be defined (field 'Table'). Enter the name of the macro, i.e. the name that is to appear in menu 'Macros', in field 'Name'. The definition itself should go in field 'Content'. Field 'Descr.' ('Description') offers additional space to comment on the macro.



➲ In the field 'Group(s)' you can enter user names or group user names. By this you specify which users may work with which SQL definition. Users with SYSOP privileges see all definitions.

The macro definition itself is made of lines which either consist of a macro command (in square brackets) or which are composed of normal characters to be written in the keyboard buffer. In the following a description of these two sorts of macro lines.

35.8.1 Macro commands

Important: Always put macro commands in between square brackets!

[DoEvents]: Work through everything before going further; **[PrgMain]**: Start Archivista 2009/I; **[Prg:Header]**: Start other program; **[Fieldname:Name]**: Corresponding field value will be given out via keyboard buffer; **[ClipboardClear]**: Clipboard contents are deleted; **[ClipboardField:Name]**: Value of field 'Name' is copied onto clipboard; **[ClipboardText:Text]**: Text is copied onto clipboard; **[ClipboardRemoveNonsens]**: Repeated spaces and returns are removed from clipboard; **[ClipboardFileLoad:Name]**: Load ANSI file 'Name' as text onto clipboard; **[ClipboardFileLoadRtfTxt:Name]**: Load RTF file 'Name' as text onto clipboard; **[ClipboardFileSave:Name]**: Save content of clipboard in file 'Name'; **[PageSave:Name]**: Save current page of document under 'Name' as graphics file; **[PageImport:Name]**: Import file 'Name' as page in current document (only 'Page view'); **[Wait:Sekunden]**: Wait mentioned number of seconds; **[Wait:Clipboard]**: Clipboard is tested every second (until there are data); **[SingleStepOn]**: Single step mode is activated; **[SingleStepOff]**: Single step mode is deactivated; **[Break]**: Macro is interrupted; **[End]**: Macro is finished.

35.8.2 Normal characters

If you want your macro to enter normal characters in any field you write straightforward characters on a separate line (Example: 'Hello'). Special characters can be entered as ASCII values but must be put in between asterisks (e.g. *32* for a space). In the case of some **special keys** and the function keys you have to take the following conversion rules as your guide:

{**BACKSPACE**}, {**BS**} or {**BKSP**} = Backspace key; {**BREAK**} = Break key; {**CAPSLOCK**} = Caps Lock key; {**DELETE**} or {**DEL**} = Delete key; {**DOWN**} = Down arrow key; {**END**} = End key; {**ENTER**} = Enter key; {**ESC**} = Escape key; {**HELP**} = Help key; {**HOME**} = Home key; {**INSERT**} = Insert key; {**LEFT**} = Left arrow key; {**NUMLOCK**} = Numerical block; {**PGDN**} = PageDown key; {**PGUP**} = PageUp key; {**PRTSC**} = Print screen key; {**RIGHT**} = Right arrow key; {**SCROLLLOCK**} = ScrollLock key; {**TAB**} = Tab key; {**UP**} = Up arrow key.

The **Function keys** are to be entered in the following manner: {**F1**}, {**F2**}, {**F3**}, {**F4**}, {**F5**}, {**F6**}, {**F7**}, {**F8**}, {**F9**}, {**F10**}, {**F11**}, {**F12**}. **Key combinations** with **Shift**-, **Ctrl**- and/or **Alt key** are realized in the following fashion:

^ = Control key (Example: `^d` = Control combined with key 'd')

+ = Shift key (Example: `+{TAB}` = Shift comb. with tab key)

% = Alt key (Example: `%f` = Alt combined with key 'f')

36 Define Fields and Masks

Tutorial on Define Fields and Masks (movie 2min) *5.5.3* **Tutorial on Additional Fields (movie 2min)** *5.6.1*

36.1 Additional fields

Table 'Archive' of Archivista 2009/I consists basically of a number of fields. These are stored in the database. After creating a new database the fields 'Title', 'KeyWords' and 'Persons' as well as some status fields like 'Document', 'Pages' etc. are at your disposal. While the status fields cannot be altered and the field 'Title' cannot be modified or deleted either (but it can be blanked out, please see 35.1.4), the fields 'KeyWords' and 'Persons' are optional. I.e. they do not have to appear in the database and can be removed. What is more, you may create additional fields.

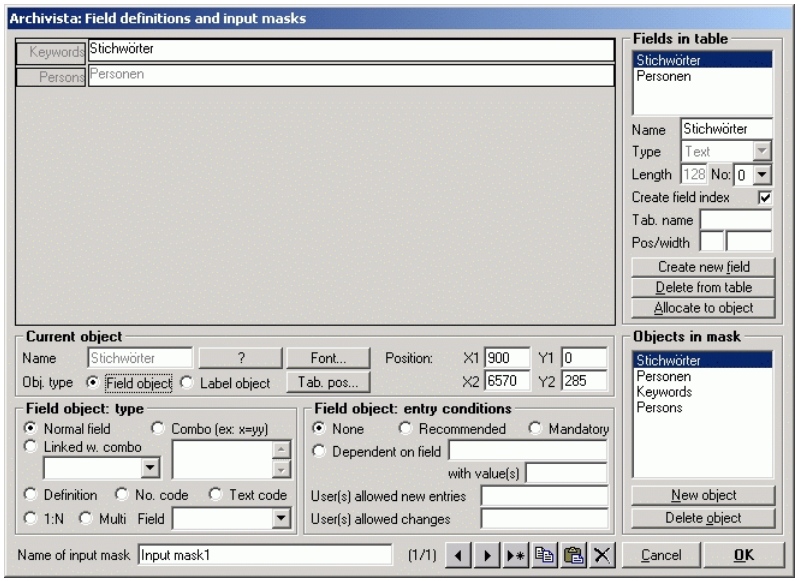
The following subchapters give you information on the functions that you need to add, modify and delete fields. Also the positioning of new fields in the mask of the detail view area is explained below.

➲ There are a variety of possibilities to create input masks in Archivista 2009/I. We recommend that you practise with a demo database first before you change existing masks in a live database.

36.2 The form 'Field definitions and input masks'

After running the submenu item 'Define fields' under 'Parameters' you get a form consisting of the following areas:

- Screen mask (top left)
- Fields in table (top right)
- Objects in mask (bottom right)
- Current object and options for field object (center left)
- Field object: type (bottom left)
- Field object: entry conditions (bottom center)
- Name of input mask (at the very bottom, left, Archivista Enterprise only)



The **screen mask** (top left) corresponds to the modifiable part of the detail view of table 'Archive' and the dialog box for searching. Here the **fields of table 'Archive'** (top right) can be displayed as objects, i.e. as cells and labels.

To create cells and labels you work with **objects** (bottom right). There is no absolute need for objects but they are often useful. An

object is either a field object or a label object. The latter serves to give an identifying name to the field object.

You see the difference between a field object and a label object by their background in the screen mask. A white object is a field object, a gray one a label object. The areas **'Current object'**, **'Field object: type'** and **'Field object: entry conditions'** serve to add particular functions to the fields.

At the very bottom of this form you see that you can create and administrate more than one input mask (Archivista Enterprise only). In fact, the system allows you to handle up to **16 field definitions**, which you can allocate to different users or groups.

➲ Fields store information in table 'Archive', objects represent fields and text (labels) in the mask.

36.3 Working with fields

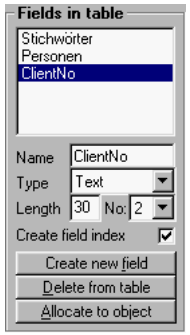
Under **'Fields in table'** you define the settings which have a direct impact on table 'Archive'. You can add individual fields to the database, you can delete these fields later and you can also create a field object and allocate it to an existing field by means of the button 'Allocate to object'. In addition you can specify what the field will look like in the table of the Archivista main view by setting table name (Tab. name), position and width.

36.3.1 Create new field

Let us first consider the case of adding a new field. Click the button **'Create new field'.** Name your field and confirm your entry with 'OK': the field is being added. At the same time you will be asked if a field object should be created, too. We recommend that you confirm this since a new field is usually meant to appear in the mask.

36.3.2 Adjust field settings

After creating a new field you should adjust the field settings. Some of these settings cannot be modified later on.



The following settings are possible:

Name: Name of the field. Please note that the name must be unique, must not contain any spaces nor begin with a number. The name of the field can be modified as long as you are still in the 'Define fields...' mode. Should you have left the form and want to modify the name, you need to do that directly on the MySQL console.

Type: We differentiate between the following field types: Text (1-250 characters), Number (*int,* i.e. integer), Number (*double,* meaning double-precision floating point to express fractions), Date and YesNo field. The field type cannot be modified in later sessions.

Length: When the field type is 'Text' you may vary the length of the text by entering a value here. Cannot be modified in later sessions.

No: Designates the position of the field within the table. Position '0' means that the field is inserted as the first field in the table (directly behind 'Pages'). In general we can say that the higher the position number the further back the positioning of the field. Can be altered in later sessions.

Create field index: With this option you determine if an index (auxiliary sorting) is created along with the field. This setting is often highly useful. Can be modified in later sessions.

Tab. name: Here you can specify the field name that is to appear in the column title of the main view of Archivista (table 'Archive').

Pos/width: In the field to the left you put the position that it should take in the main view table. Use any integer larger than '0.' In the field to the right you specify the column width the field should take in the main view table. One screen pixel corresponds to 15 units, i.e. if you enter 900, the field will be 150 pixels wide.

➲ The value '-1' means that the field in question is not to be displayed in the main view table.

36.3.3 Delete from table

By clicking this button you delete a field definitively from the table. Please be aware that **any information that is stored in this field will be irrevocably erased in all documents.** Use this command only if you are absolutely sure that the database must be rid of the field in question.

36.3.4 Allocate to object

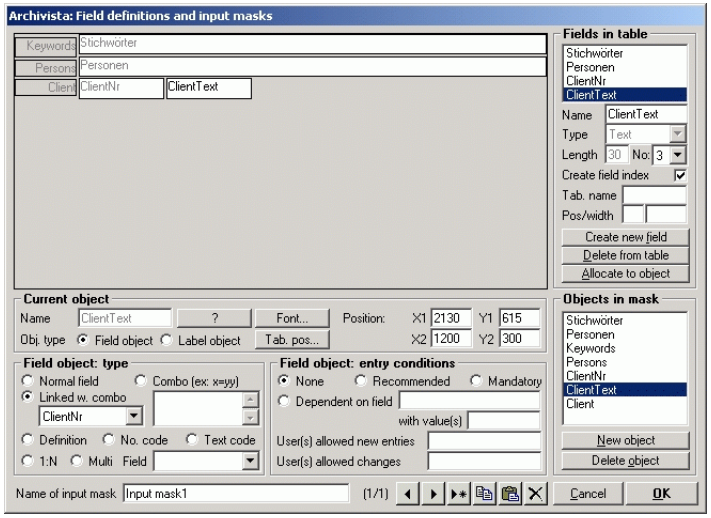
By clicking the button **'Allocate to object'** you create a link between a field from the table and a previously created label object which is subsequently turned into a field object. You create a label object by means of the button 'New object' under 'Objects in mask'.

It is not possible to allocate a field to a **field object** in the mask with 'Allocate to object'. You need in all cases a **label object.**

36.4 First steps with objects

In contrast to the fields, which are used to extend table 'Archive' (i.e. the database itself), the objects serve to shape a user-friendly screen mask for comfortable data entry.

In order that you may understand better how objects work we shall use an example. Take a look at the illustration below:

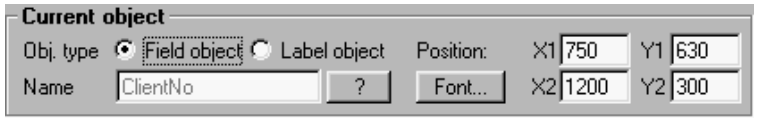


In the upper left corner you have the standard screen mask which has been extended by the fields 'ClientNo' and 'ClientText'. The field objects were created along with the creation of the fields themselves (the corresponding question was confirmed). The label object 'Client' was created manually by clicking the command **'New object'** in order to give the field objects a proper name.

➲ You may, of course, have two label objects, one for each field object. Typically you might use one line per field and call the label objects the same as the field objects, namely 'ClientNo' and 'ClientText', respectively.

Let us go back to our example. We have three additional objects as you can see in the list under 'Objects in mask'. The field object active at the moment is 'ClientNo'.

Since the object 'ClientNo' is the currently active one, the settings to the left pertain to it. To position the field object as illustrated above you must enter the following values:



You may enter the **positioning values** by way of the keyboard or you may move the active object in the mask directly with the **mouse.** If you choose the latter method you activate the object with a mouseclick. Let go of the mouse button and move the object to the place where you want it.

➲ The button 'Font' enables you to allocate a font or another attribute to a object. The button **'Tab. pos...'** ensures that in the input mask or search mask you can change the order of jumping from one field to the next by means of the tab key.

36.5 Saving the form

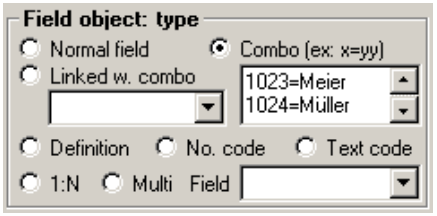
Click **'OK'** to save the form. The changes made are saved in the database table 'Archive.' The result is now visible both in the table (upper part of screen) and in the mask of the detail view (lower left).

➲ By clicking 'Cancel' before the changes have been saved all alterations are ignored. This also happens if you go back to the main view of Archivista without clicking the 'OK' button.

36.6 Field object: type

Tutorial on Field Object: type (movie 2min) *5.5.4*

Under 'Field objects' you can specify the type of field you want. Let us have a look at the settings for the previously created field 'No':



In the example above we have decided for a linked combo field and have added a couple of values. In the following the different field types are discussed.

36.6.1 Normal field

Usually **'Normal field object'** is activated in order that the field appears as a simple cell in which to enter information.

36.6.2 Combo box

Tutorial on Combo Box (movie 2min) *5.5.5*

By activating **'Combo (ex: x=yy)'** you have the possibility to create a combo box. Depending on which further entries you will make, Archivista knows three different combo fields.

36.6.2.1 Combo box with memory function

The **combo box with memory function** retains the values that were entered in the recent past. When creating a new document

the last values of the field in question are at your disposal and need not be entered afresh but can be chosen from the list. To get a combo box with memory function the area below the option 'Combo' must remain empty.

36.6.2.2 Combo box with the possibility to choose

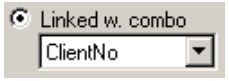
If you enter the same information in the same fields again and again we recommend that you work with the **combo box that lets you choose from fixed values.** Enter the desired values in the area below 'Combo.' Later, when entering information in the main view you can either enter the first few letters and the rest will be completed automatically or you click first on the list and secondly on the value of your choice.

36.6.2.3 Linked combo box

The benefit of the option **Linked with combo** lies in the fact that you have two fields which belong together. You want the integrity of both fields to be intact at all times. As an example may serve that you want to allocate every document to a client whereby a client consists of a client name and a unique number.

How are you to proceed? Activate the field object that is to house the number. In our example it is 'ClientNo.' Enter the linked values in the area below 'Combo (ex: x=yy)' by sticking to the following format:

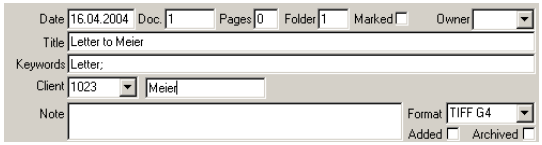
Use the sign '=' as a separator (e.g. '1023=Meier'). Now we need a second field which we link to the previously created combo field. We create the new field 'ClientText' and we link it to the combo field (notice that the option **'Linked w. combo'** is activated):



With this the two fields 'ClientText' and 'ClientNo' are linked. The entries to the left of '=' will appear in the field 'ClientNo', the entries to the right of '=' will appear in the combo box 'ClientText'.

Later, when you create documents and enter information in the main view you may pick '1023' from the list in 'ClientNo' and make 'ClientText' jump automatically to the corresponding 'Meier'. You may, of course, also choose the opposite procedure.

In the end you should get the following screen mask as part of the detail view in table 'Archive' main view. (The field 'Persons' has been eliminated since we are working with 'Clients' now.)



36.6.3 Definition

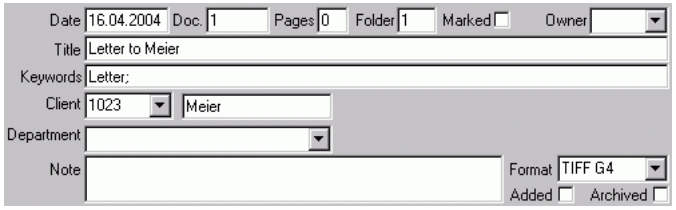
Tutorial on Definition (movie 2min) *5.5.5*

The combo fields described above, whether linked or not, are a bit limited in their use. You have to click on the field to choose a value. The field types described in the following paragraphs offer more convenience because you can add values in the main view without having to take recourse to the 'Define fields' form and because there is **Table 'Field lists'** if you nevertheless want to maintain the lists comfortably.

Let us take a look at the simplest case, namely the field type **'Definition'.** This field type is used whenever you want a simple combo box both for input and search.

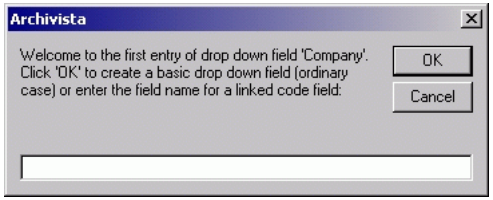
We explain its functionality by considering the following example. Run the function 'Define field ...' and create a new field with the name 'Department.'

Go to 'Field object: type' and click on **'Definition'**. Save the field definition by clicking 'OK.' You should now see the following input mask:



When you click the field 'Department' you see that the combo box is as yet empty (no values have been defined):

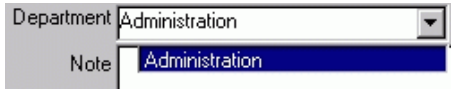
Now add the value 'Administration' and press the tab key: the following message pops up (only when you define the first entry).


Archivista
Welcome to the first entry of drop down field 'Company'. Click 'OK' to create a basic drop down field (ordinary case) or enter the field name for a linked code field:
OK
Cancel

Please click **'OK'** since we want to create only a basic drop down field. A further message appears:

Archivista
The definition 'Administration' is not yet part of the standard values (table 'Field lists'). Would you like to create a new entry?
Note: If you press 'No' the value will be added to the active record but will not be saved as standard value.
Ja
Nein
Abbrechen

This kind of confirmation screen comes up when there is no entry yet in table 'Field list' for a particular field (here 'Department'). Click **'Yes'** and the value 'Administration' will be definitively added to table 'Field lists.' If you go back to the field 'Department' now by means of 'Shift+Tab,' you see that the value 'Administration' has been added:



Let us add the values 'Production' and 'Sales' to our field 'Department.' Now you should see a screen roughly like the following:

➲ Please note that when choosing a value from your drop down list it is sufficient to enter the beginning of the word up to the first distinctive character. The value you want is selected and by clicking 'Enter' you transfer it to the field.

36.6.4 Linked definition fields (No. code, Text code)

Similar to when you work with linked combo boxes you can also link fields when you work with the field type 'Definition' and the resulting table 'Field lists.' We distinguish 'Definition,' 'No. code' and 'Text code.'

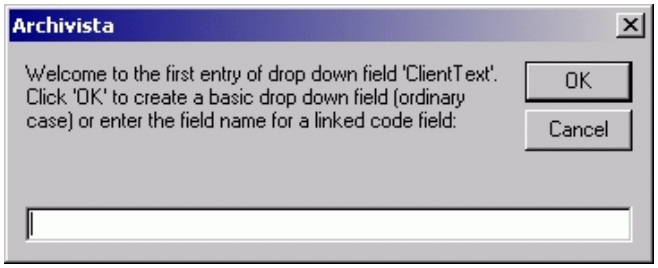
The field specified as 'Definition' will always contain the text in full length while the other fields, 'No. code' and 'Text code' give a home to the appropriate number or abbreviation. Again, we best give an example to illustrate the use of these fields. Run the function 'Define fields ...' and create two new fields. Name the first field 'ClientText' and choose 'Definition' under 'Field object: type.' Name the second field 'ClientNo' and choose 'No. code' as type.

➲ If you have already worked through the example of the linked combo box and if the fields 'ClientText' and 'ClientNo' exist al-

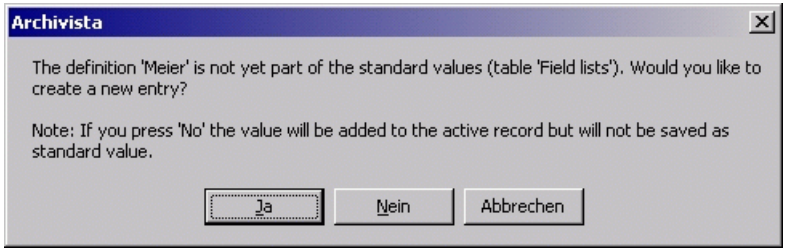
ready, then simply change the field type to 'Definition' and 'No. code,' respectively.

Save your changes by clicking 'OK.' You should see roughly the following input mask:

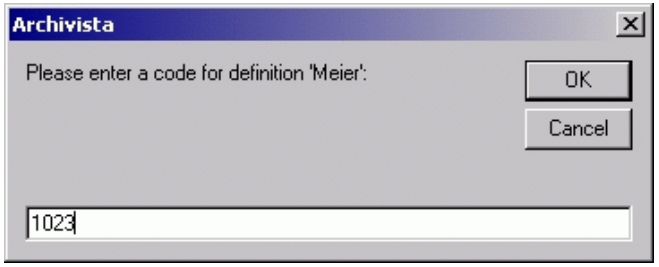
At first sight you do not see a difference: the fields look like two ordinary drop down fields or indeed like the linked combo box we created earlier. But now, click the field 'ClientText' and enter the value 'Meier.' Press the tab key. You will get the following screen:


Archivista
Welcome to the first entry of drop down field 'ClientText'. Click 'OK' to create a basic drop down field (ordinary case) or enter the field name for a linked code field:
OK
Cancel

As we are keen to create a linked field of the type 'Definition' now, please enter the name of the field to be linked: in our case it is 'ClientNo.' After clicking 'OK' you will see the following message that again needs to be confirmed.


Archivista
The definition 'Meier' is not yet part of the standard values (table 'Field lists'). Would you like to create a new entry?
Note: If you press 'No' the value will be added to the active record but will not be saved as standard value.
Ja
Nein
Abbrechen

In order to finish the definition of our entry we need to add a client number (field 'ClientNr'). Enter the value '1023,' for example, and click 'OK.'



The first predefined value has now been created. Do add a couple of more values also for this example. We suggest that you use 'Müller' with '1024' and 'Bucher' with '1025.' To test the link choose '1025' in the number field and you will notice that the field 'ClientText' will automatically show the corresponding 'Bucher' (see the illustration below):

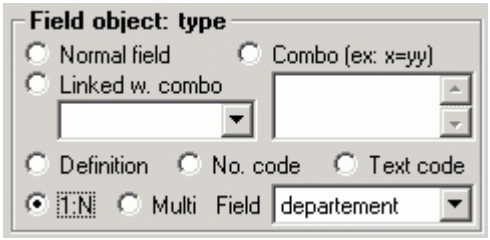
By adding more values you will realize that both creating new values and choosing from existing ones can be done fast and easily.

➲ When using the field type **'Text code'** you can work with abbreviations instead of numbers. For example, you can use 'me,' 'mü' and 'bu' for 'Meier,' 'Müller' and 'Bucher' if you like. It should be noticed that when first creating this kind of field you choose 'Text' under 'Fields in table' whereas with a 'No. code' field you choose 'Number.'

36.6.5 1:N (Hierarchical)

The field type '1:N' enables you to build drop down lists which depend hierarchically on specific entries in other drop down lists, i.e. with this field type you can compile veritable keyword tree structures. Let us make an example. We create an additional field which is dependent on the field 'Department': let us call it

'Region.' And now the essential option: we activate **Field type '1:N'** and we choose 'Department' in the next but one field to the right. We have thus specified a 1:n relationship between the fields 'Department' and 'Region.'



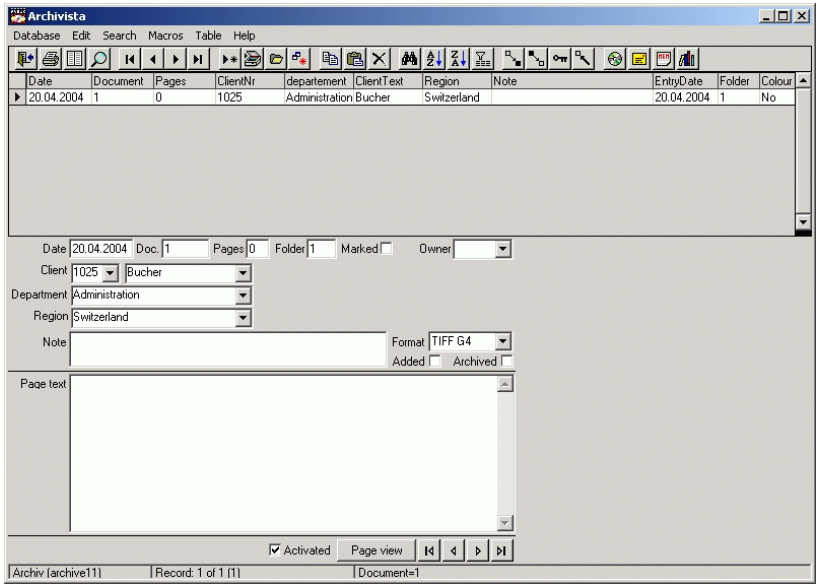
By saving the changed field definition you are taken back to the main view. Go to the detail view or input mask in the bottom left corner and under 'Department' choose 'Administration.' Now go to the new field 'Region' and create entries like 'France,' 'Benelux' and 'Switzerland,' for example. After having created these entries choose 'Switzerland.'

Now create a new document and under 'Department' choose 'Sales.' When you go to the field 'Region' you will notice that the previously created entries 'France,' 'Benelux' and 'Switzerland' do not appear: hierarchically they belong to the department 'Administration.' For the department 'Sales' you could now compile the values 'Switzerland,' 'Germany' and 'Italy,' for example, and select 'Italy' for correct keywording if you like.

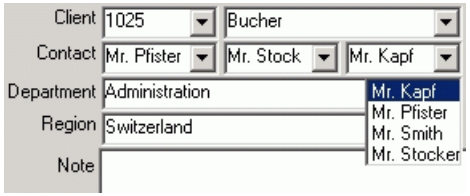
➲ 1:N fields can extend over more than two hierarchical tiers. You could, for example, define a 1:n relationship between the field 'Region' and a new field 'Contact.' By doing this you would have created a third tier. Similarly, you could make several 1:n relationships for one field type 'Definition.' You must simply see to it that the field type 'Definition' is top of the hierarchy.

36.6.6 Multi field

The field type 'Multi' enables you to add more than one keyword even if you work with the controlled vocabulary of drop down lists. You can, for instance, select several contact persons per document if you work with this field type. The following example may serve to illustrate this. Run the 'Define fields ...' function and create three fields with the names 'Contact,' Contact1' and 'Contact2.' Make the field 'Contact' the 'Definition' and choose 'Multi' for the fields 'Contact1' and 'Contact2' connecting them with 'Contact' by selecting 'Contact' under 'Field.'

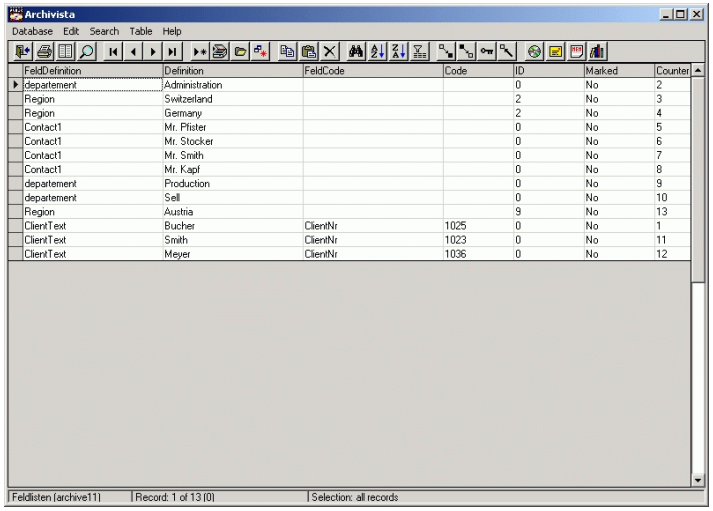

Archivista
Database Edit Search Macros Table Help
Date Document Pages ClientNr departement ClientText Region Note EntryDate Folder Colour
20.04.2004 1 0 1025 Administration Bucher Switzerland 20.04.2004 1 No
Date 20.04.2004 Doc. 1 Pages 0 Folder 1 Marked Owner
Client 1025 Bucher
Department Administration
Region Switzerland
Note
Format TIFF G4
Added Archived
Page text
Activated Page view
Archiv (archive11) Record: 1 of 1 (1) Document=1

Now you can allocate more than one contact person to a document. All values predefined for 'Contact' appear also in the fields 'Contact1' and 'Contact2'. See the illustration below:



36.6.7 Table 'Field lists'

After having discussed all possible field types the question might come up how exactly the predefined values are administrated. The answer is to be found in the auxiliary table 'Field lists.' If you have gone through all the examples you should see roughly the following entries:


Archivista
Database Edit Search Table Help
FeldDefinition Definition FeldCode Code ID Marked Counter
departement Administration 0 No 2
Region Switzerland 2 No 3
Region Germany 2 No 4
Contact1 Mr. Pfister 0 No 5
Contact1 Mr. Stocker 0 No 6
Contact1 Mr. Smith 0 No 7
Contact1 Mr. Kapf 0 No 8
departement Production 0 No 9
departement Sell 0 No 10
Region Austria 9 No 13
ClientText Bucher ClientNr 1025 0 No 1
ClientText Smith ClientNr 1023 0 No 11
ClientText Meyer ClientNr 1036 0 No 12
Feldlisten (archive11)
Record: 1 of 13 (0)
Selection: all records

➲ You need 'SYSOP' privileges to be able to access this table by means of menu 'Table' or 'Ctrl+Shift+F11.'

The table 'Field lists' is rather self-explanatory - maybe apart from the field 'ID,' in which it is coded how hierarchical fields are linked. Take a look at the entries where it says 'Region' under 'FieldDefinition.' The first two entries show a '2' under 'ID': this means that these fields are linked with the field possessing the counter '2.' It is 'Definition' 'Administration' ('FieldDefinition' 'Department'). In the same way ID '9' refers to the entry with the counter '9,' namely the department 'Production.'

➲ In principle you can change the entries in table 'Field lists' directly by adding or deleting them. You can edit existing entries by clicking on a field and pressing the 'Enter' key. However, working with the input mask is more comfortable. We suggest that you work with table 'Field lists' mostly when you want to delete false values or when you want to make a batch of corrections. Please note that when you delete values from table 'Field lists' they will not occur anymore in the drop down lists but they remain in table 'Archive,' i.e. values previously selected for the keywording of documents remain allocated to these documents.

36.6.7.1 Editing field lists by means of the search mask

In menu 'Search' of the Main view the first submenu to be found is 'Documents (fields)...(F5)'. When you run this command or press 'F5' you can call up the search mask.

With this mask you can edit or delete fields from field lists. While – as the name suggests – the search mask is primarily a tool to find records with, it fulfills a double function in connection with field lists.

When you have to go through complicated editing or deleting sessions working with the search mask is superior to working with table 'field lists'. Please note that you can solely edit or delete fields. You cannot add any entries.

The procedure is the following:

- Pressing 'F5' takes you to the search mask
- Click the field in question so that the pull down menu opens and move the cursor over the entry you want to change or delete

- **Double click** the entry
- A dialog box pops up. Click 'Yes' and the entry is removed from the list, click 'No' and you can edit the entry. By clicking 'Cancel' you get back to the search mask.

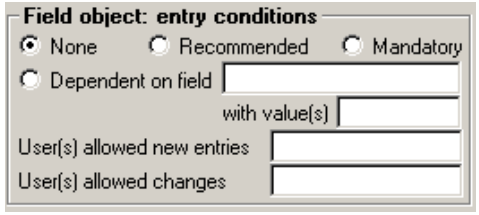
When editing hierarchical fields it must be noted that main entries can only be removed if all corresponding entries from the subvalue list have been deleted. This is a security measure to prevent that entire filing trees are removed.

36.7 Entry conditions

Tutorial on Entry Conditions (movie 2min) *5.5.6*

With 'Entry conditions' you can set limits to what you want to accept as input. Additionally, you can specify that not every user may be allowed to change field content, or only if there is a dependency from another field.

Let us take a look at the options under 'Entry conditions':



36.7.1 None

Whenever input should be without restrictions of any kind choose **'None'**. All entries will be accepted as valid.

36.7.2 Recommended

The option **'Recommended'** means that a field should not remain empty. You can enter all sorts of values but if eventually you leave the field empty, a message will pop up inviting you to fill in a value.

36.7.3 Mandatory

When **'Mandatory'** is activated users cannot leave the field without having entered a value.

➲ Please note that often when you first create a document there are fields that are still empty. An empty field may remain empty until the user jumps into it. If it should be made impossible to have empty fields on creation of a document you must change the default value of a field or force the user to make an entry with the help of the workflow module of Archivista Enterprise.

36.7.4 Dependent on field with value(s)

This is the option you choose if you want that a field accepts any values only on the condition that there are specific values in another field.

Under **'Dependent on field'** you specify the name of that other field. Under **'with value(s)'** you put the specific value the other field must contain to enable you to make an entry in this field. The following example may illustrate the use of this option: You have a field 'SubjectArea' and one called 'ContractNumber.' The users should only be enabled to enter a contract number if the value 'Contract' is selected for 'SubjectArea' since a contract number without contract does not make any sense.

36.7.5 User(s) allowed new entries

Here you can specify users or groups of users that are allowed to make new entries (for example for drop down lists). The other

users may also add keywords to new documents but they may only enter the values used before (i.e. those for which there is already a record in table 'Field lists'). The primary aim of this option is the prevention of typographical or orthographical errors.

36.7.6 Users allowed changes

Only the (group) users specified here are entitled to make changes to the content of this field. The aim is to give 'read only' access on the level of the field. You exclude a (group) user from the right to change something but at least they may see what the field content is.

➲ Please appreciate the difference between this option and the complete masking out of a field (both in the input mask and the table). With the option discussed above you prevent users from changing a field. When a field is masked out completely the users cannot see that the field is there at all.

36.8 Different input masks

With Archivista Enterprise different users can work with altogether different input masks. You find this option at the very bottom:

Under 'Name of input mask' you can specify a name. To the right of this field you find icons with which to administrate the at present maximum number of 16 definitions. Please note that you can copy entire definitions (commands copy and paste) and that you can delete them too.

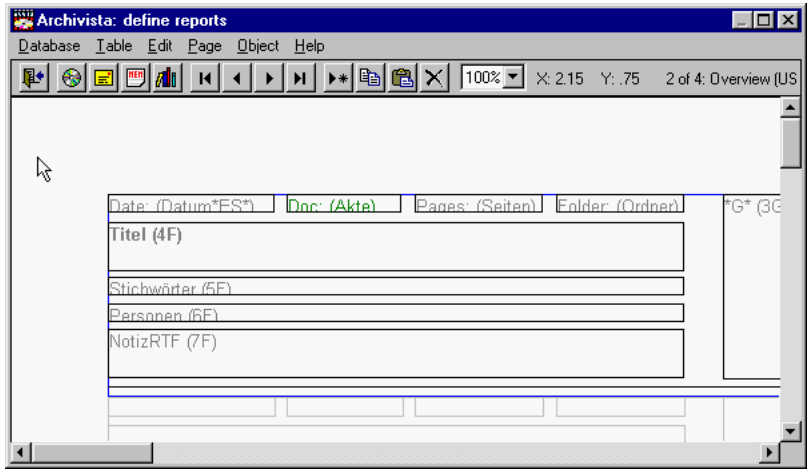
➲ After you have created a new input mask you must allocate it to the users concerned. You do this by running the function 'User administration ...' in menu 'Database' / 'Parameters.' Please note too that only the first definition allows you to add or delete fields. And you can only do this if you have not yet made changes to the other definitions. Should you want to add new fields or delete fields after having edited the secondary input masks, you must save your changes first and run 'Define fields ...' afresh.

37 Define Print Reports

Tutorial on Define Print Reports (movie 2min) *5.4.4*

37.1 Function

The report module helps you to create so-called print reports for the main tables. These are definitions to print your data. When you run this submenu item you get to the **'report module':**



The report module always shows one page in a page preview in order that you may draw your report, i.e. place **'print objects'** on the page.

➲ To access the report mode you need systems operator privileges. The workings of the report mode are not described in as much detail as other functions in this manual. First, there are templates to work with, secondly only able users should try their hand at print reports.

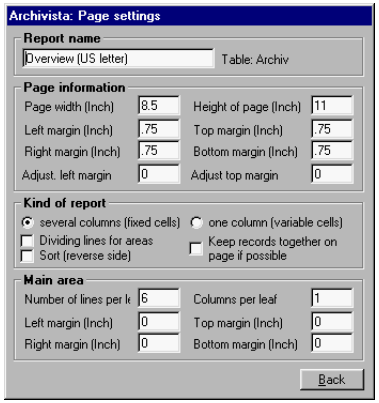
In principle the following steps are needed to define a print report (printer definition):

37.2 Create new report

First choose the desired table for which you want to define a print report. Then go to menu **'Edit'** and menu item **'New report'.** By default an A4 page portrait fashion is created.

37.3 Page setup

Go to menu **'Page'** and enter page information that applies to the whole page.



The following settings have to be specified.

37.3.1 Report name and group(s)

Define the name which later will appear in the listing in menu 'Print' (menu 'Database').

➲ Each report should have its own, non-equivocal name because otherwise the listing might show several entries of the same name.

Specify in field 'Group(s)' which user(s) may use this particular report. By default this field is empty, i.e. by default all users may print the report.

➲ If you prevent certain users from seeing certain fields you should also make sure that they cannot print fields by way of the print reports.

37.3.2 Page information

Here you define the **measurements** of your report page and the **margins** as well as adjustment values for an optimized appearance of your data on paper.

➲ Please note that you do not have to enter information on whether the report is to be printed in portrait or landscape style. As soon as the page width is larger than the height of the page Archivista assumes that the page is to be printed in landscape fashion.

The **page measurements** are **not transmitted to the printer.** The advantage is that you can use all paper sizes. You must make sure, though, that the set measurements do not exceed the physical size of the sheets. With some printers it may be necessary to move the whole page slightly to the left or right or upwards or downwards. The fields **'Adjust left margin'** and **'Adjust right margin'** may be used for this.

37.3.3 Kind of report

Choose between **reports with several columns** and fixed cells and **reports with one column** and variable cells.

The option **'Sort (reverse side)'** helps you to print index cards with text also on their back. What it does is it exchanges columns when printing. In other words: the column on the very left is printed on the very right and vice versa, the second column on the left is printed as second on the right and vice versa, and so on. What is more, you can specify here whether **dividing lines** should be printed between the different areas (cells).

37.3.4 Main area

Here the individual **cells** of the print reports are defined. By cells we mean the area which is reserved for one record. The number of cells is the result of 'number of rows' times 'columns per sheet'.

When you go for **variable cells** only one column is possible. When entering the number of rows per sheet please take heed that the resulting height is minimal. If a line takes more room later (because of an object with variable height), the height of the row is adjusted automatically. In other words: row height is variable.

➲ The first cell of the report module (painted blue on the screen) is the one on which you later place your objects. What you define for the first cell, i.e. for the first record, applies to all records afterwards.

37.4 Defining objects

By objects we mean in principle a fill-in to print certain data. Per report you can place 50 objects as a maximum. Do not confuse objects with reports. In contrast to the reports the objects are not created via menu 'Edit' but via menu **'Object'**. See menu items **'Add', 'Duplicate'** and **'Remove'.** A print report must at least contain one object.

You place the object manually on the page by left-clicking it or by choosing it via menu item **'Select'**. The activated object has a colored frame. You determine size and width of the active object with your mouse (grasp the object at the bottom right by pressing your left mouse button, keep the button pressed and move your mouse to enlarge or diminish the object). You find further settings in menu 'Object'. You can also use the shortcut menu (right mouse click) to make the same menu items appear. In the following a description of the additional settings.

37.4.1 Printing position

With this item you determine the order in which the objects are printed. What is printed last is uppermost on paper. Do bear in mind that graphic objects usually make text objects unreadable when printed on top of them.

37.4.2 Kind of object

There are five kinds of objects:

- **Field objects** contain a field from the corresponding table.
- **Text objects** contain text (but also the current page number, for example)
- **Combined objects** consist of fields and text bits (see below)
- **Graphic objects** stand for internal (from table 'Archive') or external graphics
- **Frame/Line** objects serve to print frames and lines

➲ Before you specify the settings of an object you must determine the kind of object you want!

37.4.3 Special entries in objects

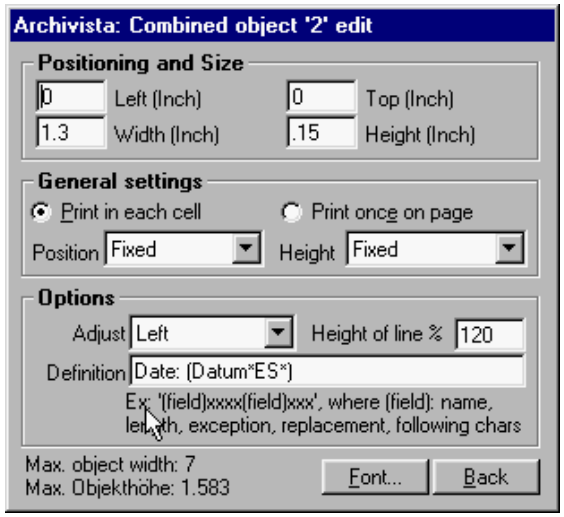
Field objects: Each date field appears four times, i.e. once for each display possibility (see combined objects). Memo and RTF fields (see below) appear twice as well as the Yes/No fields (with addition of *E*='Yes/No' and *D*='Ja/Nein').

Text and Combined objects possess the following variables: *P* = current page number, *S*=page total, *DS* = printing date 'dd.mm.yy', *DL* = 'dd.mm.yyyy', *ES* = 'mm/dd/yy', *EL* = 'mm/dd/yyyy', *L* = counter.

Graphic objects contain either the name of the graphic that is to be tied in or – if the first page of the document record is to be incorporated – the following abbreviations: *G* = screen copy quality (low resolution), *GD* = printing quality (high resolution).

37.4.4 Settings

Under **'Settings'** we find 'Positioning and size', 'General settings' and 'Options'.



Positioning and size: Enter the position of the object. When you go below minimum or above certain maximum values an error message ensues.

General settings: If you activate 'print in each cell' then the object is printed as many times as there are cells (number of rows x number of columns). If you activate 'Print once on page', then an object is printed only once (e.g. for headers or footers).

Under '**Position**' the following settings are possible:

Fixed: the object has a fixed position

Variable: the position of the objects is calculated on the basis of the variable heights of the previously printed fields. Please note

how Archivista proceeds: only objects just above the mentioned object and within the same width are taken into consideration when the variable position is calculated.

Below: positioning takes place in accordance with the variable object height at the bottom of the cell. This helps you, for example, to print a line at the bottom of a cell.

Also under **Height** you can choose between **Fixed** and **Variable.** 'Fixed' is the default, 'Variable' can only be selected for a text object, a field object or a combined object. If an object of variable height is activated, the height is calculated on the basis of the content of the record. Cells below take their position accordingly.

Please note that with the **variable (single column) print reports** the height of the cells (and of the rows) is calculated in dependence of the variable heights of the objects. The maximum height of a variable object is therefore only limited by the page measurements. In the case of the fixed reports of several columns the height of the cells is at the same time the maximum height of the objects!!!

Options: Here the entries depend on the kind of object you want to place on your report. The respective fields should be self-explanatory with the exception of the combined object which offers you comprehensive possibilities that are described in the following passage.

37.4.4.1 Combined objects

Combined objects are fields that contain several text bits and parts of fields. The following format must be adhered to:

`xxxxx(field)xxxx(field)xx*ASCII*xxxx'`

- **'xxxxx':** means normal text (without ASCII special characters), e.g. **'Date:'**
- **(field):** means a **value from a field** and consists itself of several parts, namely, **'name, length, exception, replacement, subsequent character'** (the individual parts are to be divided by commas).
  - **name** here you enter the field name
  - **length** here the maximum length of the field value
  - **exception** her the value which should not be printed
  - **replacement** here you enter the value which should be printed instead
  - **subsequent character** further text characters are possible. An example: '(country code,5,CH,,-) means that the field 'country code' is to be printed, that a maximum of 5 characters is admissible, that the country code 'CH' is not to be printed, that no replacement is to be introduced and that a hyphen is to be used as subsequent character (but only if a country code other than 'CH' is present).
- ***ASCII*:** enables you to print ASCII characters which cannot be represented (values 0 to 31) plus those that are used as delimiters by Archivista for the combined object (characters '(', ')' and '*'). The character sequence *13**10* results in a carriage return at the position where you put it.

37.4.5 Font

With the exception of the graphics and frame/line objects you can use an individual font definition for each object. Use the menu 'Font' to do this. By default font 'Arial, 10 point' is used, i.e. if you do not specify anything else.

37.4.6 Printing of RTF fields

RTF fields (see menu 'Format' in chapter 'Menus') can be printed either as normal text (choose the corresponding field name in the field object) or with formatting information (in this case choose the field name with the file extension 'RTF', e.g. 'NotizRTF').

Part X

Publishing

38 PublishingEdition II

38.0.7 Long-term considerations on documents

At the very beginning we should ask ourselves for which reasons we want to create an Archivista PublishingEdition. Let us assume that we want our documents to be available for a long period of time. 'Long-term' may mean lots of different things but let us think of a timespan between 20 and 50, or maybe even 100 years. This sounds like an infinite stretch of time in the world of computer technology but it is not in terms of paper documents.

Which documents will still be available at that point in time? Will we still be able to read an MS Word file properly in 20 years' time? Will the Internet still exist (admittedly a heretic question)? Where will Adobe's PDF format be? Will we still have the right to look at our own documents without a meter ticking away somewhere and intermittently making a message pop up that our account is running low?

38.0.8 The Schneider adventure

Please permit us to make a short digression here by relating an experience which had a lasting effect on our thinking. In 1983 a man walked into a shop and bought his first PC. The man was later to be managing director of Archivista GmbH and the PC was a Schneider CPC 464: it used a tape as storage medium. The salesperson in the retail shop who sold it said that the tape was a relatively unsafe medium but for short-term purposes it would be more than sufficient. Soon a floppy disk drive would come on the market and that would be extra safe and much faster too.

The storage capacity of 180 KBytes had great appeal but the Schneider computer was the only PC around with a 3" disk drive. Consequently, the shop assistant was asked if it was not a problem that Schneider had a 3" drive while all other home computers possessed either a 5.25 or a 3.5" drive? No problem, he answered, the 3" drive was very much on the upsurge, Amstrad was introducing it too.

The customer swallowed what the shop assistant said and bought the Schneider PC. Only much later did he realise that Schneider and Amstrad were the same and that his initial gut feeling proved right: no other 3" drive ever appeared on the market. The man stuck to his Schneider PC faithfully but the data of those years, i.e. 1983 to 85, could not be converted later on. It was his luck that he had printouts of all the files so that he could scan and archive them later. Without this experience Archivista might never have come into existence.

What are we to learn from this cautionary tale? In contrast to the 3" disk the tape has survived until today. And, the information could be saved because it existed as printouts. Does that mean that we have to produce printouts of all electronic files on our computer? The answer cannot be 'yes' here, also because we would not need a PublishingEdition then.

38.0.9 The XP adventure

The first version of the Publishing Edition was created within a client project in the year 2004. The task was to create a CD which could be started on a Windows PC without requiring any additional software. This solution was greatly loved by our clients right up to the moment when with Windows XP the firewall was activated with SP1 by default.

Not that our solution did not run any more, but since our solution used local port 8088, the Windows firewall blocked this partic-

ular local port. In the final analysis this meant that for most users working with our Publishing Edition went from worry-free to hiccup-laden.

We could go into arguing that it makes little sense to shut the local ports in a local environment by default when at the same time (at least at the time) it was possible to write on the hard disk via the temporary directory. However, when you want to open a self-supporting archive, then such deliberations are hardly of interest to you. You simply want to work with your archive.

From a vendor's point of view it was a deplorable situation. We did make all the right choices with Web, MySQL and Windows, did we not? Bad luck when we propose but Windows disposes. Windows is not fond of Web technologies and this is not our fault. But even this will be of little interest to you when you are trying to make archived data available.

As long-term reaction to this (and other) experiences ArchivistaBox has come into being: a solution which on the one hand keeps favouring Web technologies and on the other can be deployed as stand-alone, embedded box solution to be at all events prepared for future imponderabilities.

Until a few months ago using an ArchivistaBox required that the software be set up and the data to be archived needed to be added. Because of this, ArchivistaBox could not produce self-supporting archive CDs. It was then decided to enable ArchivistaBox to create self-supporting CDs anytime from a running system. These CDs were to serve a double purpose: for one they should be liveCDs viewable as live archives, and for another they should serve as installation CDs whereby an archive could be put back onto the server and where new data could be added.

We can only be grateful that at that time the Windows firewall shut the local ports because it gave us the possibility to make an effort towards a solution which is in every direction much more open than the previous Publishing Edition.

With the new solution (which is - it goes without saying – 100 percent OpenSource) no Window license is needed nor do we have to list any version numbers or ServicePacks as prerequisites. All we need is: a) a computer with 256 MByte RAM or b) a virtual box with 256 MByte RAM. Everything else is handled by the ArchivistaBox: and suddenly publishing archives is easy.

We are of the opinion that the Roman Numeral II in the chapter heading is justified. We wish you and us much success and pleasure with the new Publishing Edition.

38.0.10 Why source formats are not suitable

How can we ensure that our documents keep well and are safe for long periods of time? Why are source formats (e.g. MS Word files) unsuitable and why should we also convert PDF files although they carry 'portable document format' in their name? Simply because we need powerful programs to display these files. All of them contain characters, more or less exotic fonts, images, and sometimes even bits of executable code, that need to be put together on the fly and with the help of considerable processing power before we can view our file on screen.

Keeping documents can be massively simplified if we create an image file out of our source format. The image file resembles a photograph in that it freezes the look of a file at a particular moment in time. Display on screen is much simpler too. At Archivista Gmbh we speak of 'virtual printing' because converting source format files into image files entails nothing more than printing them to a file and rasterising them. Finally, they end up as image files in the Archivista database.

In the following I would like to explain the abovementioned process step by step and with the help of Archivista. It saves us the task

of having to print out our documents, file and re-scan them again. Of course, you can also publish documents that exist only as hard copies to begin with: these you scan first. However, the following tutorial is dedicated exclusively to documents already existing in digital form on your computer.

The points made in this section are valid for all archives whether they were created with the Archivista PublishingEdition or not. We should be aware of them when we later encounter problems during the import of documents. You may safely assume that if certain files give you problems when you convert them today, they would generate immensely more hassle if you tackled them in, say, ten years.

Have fun with the following tutorial!

38.0.11 Preliminary note

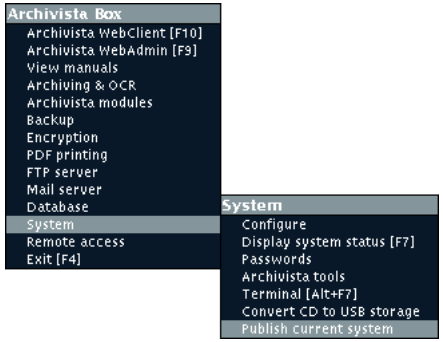
In order to publish Archivista databases you need an ArchivistaBox installation. How to install ArchivistaBox is described in 3.1. As soon as this is done you can add documents. To create a publication you leave the WebClient and run the menu item 'Publish current system' in the system menu (F12, right mouse click and then choose 'Publish current system' from menu 'system').

38.0.12 Publish current system

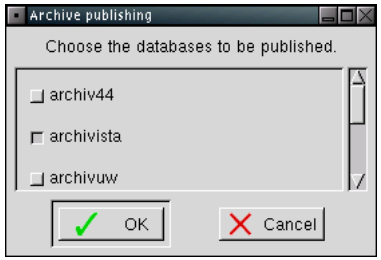
.

'Publish current system' serves to copy the entire fully functional system – i.e. the Archivista database with all documents and metadata but also the settings – to a USB or CD/DVD device. This way it is simple, for example for accountants, to take client documents along to the client site and display and search them there.

Run the command.



After password entry there appears a list from which you can choose the database which is to be published. By default 'archivista' is ticked.



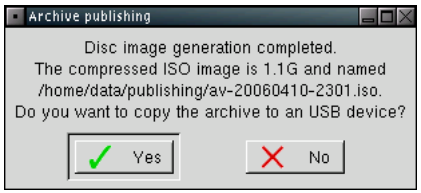
In the following you see a confirmation message mentioning the database(s) concerned.

The next screen tells you the amount of space that is needed. You can choose whether you want to have a compressed or an uncompressed system. The advantage of the compressed system is that the ISO file remains small enough so that also large databases may find room on a DVD in the end.

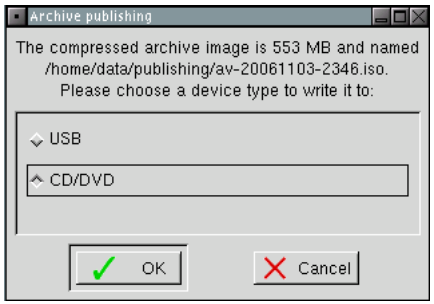
The ISO file will be created. You see by the number of megabytes mentioned at the end of the message if the process is soon coming to an end or whether you still have to wait.

➲ Please note that no user is allowed to work with the system during the time the file is put together.

As soon as the process is finished you receive a message confirming the creation of the ISO file.


Archive publishing
Disc image generation completed.
The compressed ISO image is 1.1G and named
/home/data/publishing/av-20060410-2301.iso.
Do you want to copy the archive to an USB device?
Yes
No

By the same token you will be asked if you want to copy the archive to a USB or CD/DVD device.



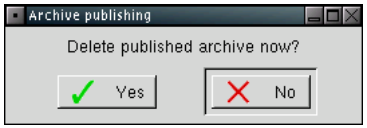
38.0.12.1 USB medium

After clicking the 'Yes' button a dialog box announces that the ArchivistaBox is waiting for a USB device.

After you have attached the USB medium the ISO file is being copied.

As soon as the copying process is finished, the following dialog box comes up.

You will also be asked whether you want to keep the ISO file. If you keep it, you can re-use it, say, for moving it to your Windows system by means of Win SCP and burning a CD or DVD. But please bear in mind that the space such a file uses is not negligible.



Now is the time to shut down the ArchivistaBox. We recommend that you test the live system on the USB device straight away. Leave the USB device attached. Use the Exit function to shut down the ArchivistaBox. Then, press the 'Start' button on the ArchivistaBox and, after a couple of seconds, press the ESC key. You are now in the start-up menu where you can choose from which hard disk (HD for hard disk) the system should boot. Choose 'HD' first and then 'USB device'. After start-up you will be seeing the live system running on the USB device.

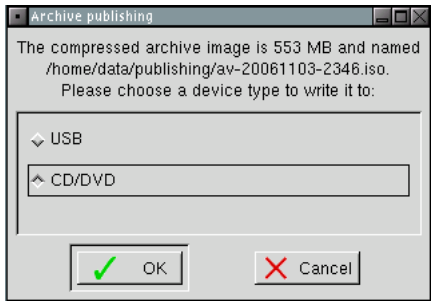
➲ Important: Do not detach the USB medium from the ArchivistaBox while the latter is still up and running. To remove the USB medium always shut down the ArchivistaBox first.

This live system on the USB device may again be installed on a (different) computer. Use the menu item 'Install current system to a hard disk partition' under 'System' for this purpose. Please note that you must make a full installation in order to have all data on the target machine eventually (option 'Reformat whole disk').

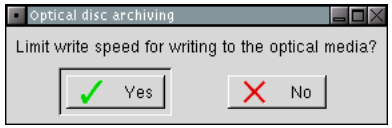
38.0.12.2 CD/DVD medium

If you know in advance that you want to burn a CD or DVD, insert the disc in the drive before you start the function 'Publish current system'. Should your ArchivistaBox come without such a drive, simply attach an external drive by means of a USB cable.

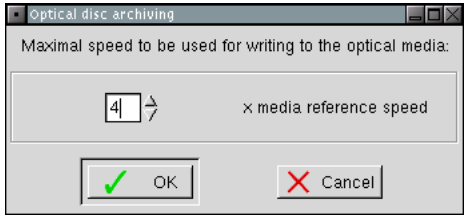
After the dialog box that asks you whether you want USB or CD/DVD (see below) ...



... you must decide whether you want to delimit writing speed. Modern CD/DVD drives can write at very high speed but the quality of the media suffers considerably. That is the reason why there is the possibility to reduce writing speed by pressing the 'Yes' button.

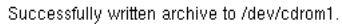


The next dialog box asks you for the speed with which the process is to be carried out.



After that the CD/DVD is being written. A message to that effect appears.

The end of the process is confirmed by a message that the archive was successfully written by means of the attached device.



➲ As described in 38.0.12.1 there appears also here a dialog box asking whether the ISO file created on the ArchivistaBox for the purpose of writing the CD/DVD can be deleted now.

➲ It is further advisable that the published system be tested straight after creation. In addition, the CD/DVD serves not only as copy of your archive but it can be used for the installation of the system on a PC. Read 38.0.12.1 for a more detailed description of the two abovementioned points.

38.0.13 Testing the medium

Media created by means of the publishing function should be tested right away to ensure that the result satisfies us.

We reboot the ArchivistaBox and make sure that it boots from the newly created medium (and not from the hard disk). We should now find those databases on the medium which before were published.

38.1 A few tips on working with documents

In the following we would like to give you a number of tips of how to get the documents into the ArchivistaBox in the first place (before publishing them in a second step).

38.1.1 Working with paper documents

Documents you receive as so-called hard copies must be scanned. Before starting this time consuming process you should make sure that no PDF file exists. As a rule you should be able to proceed much faster with electronically available documents if you possess the needed applications and the Archivista printer is set up properly on the ArchivistaBox.

300dpi yield good results when scanning. For the sake of velocity colour documents can be captured with lower resolution, however,

it must be borne in mind that the text recognition might suffer accordingly.

For larger volumes we recommend to aquire a scanner with automatic document feeder if it is possible to split the documents into individual pages. Books often defy this procedure. For books a flatbed scanner may be used or a digital camera. Please note that with digital cameras lighting is essential – artificial light might give problems – as well as a tripod to hold it firmly in position. For A4 pages a camera should possess at least a resolution of 6 mio pixels.

38.1.2 Working with electronic files

To transfer individual files to the archive we proceed in the following manner. We start up the application and open the particular file in question. Then, we print it using the printer named 'Archivista.'

➲ For information on how to set up a printer on the ArchivistaBox to print documents into the archive please see below.

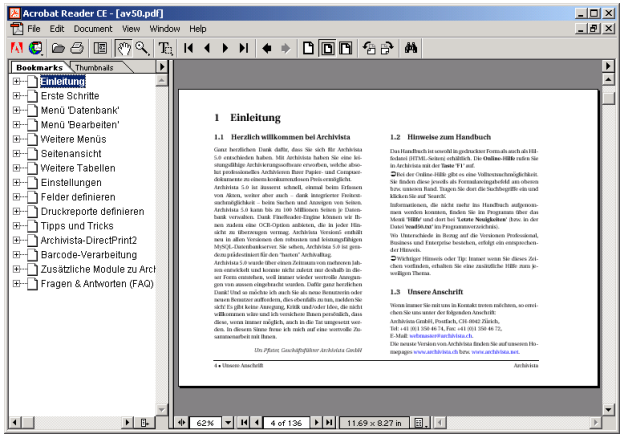
38.1.3 Working with PDF files

In the following we would like to discuss a few peculiarities of working with PDF files. To begin with let us bear in mind that the acronym 'PDF,' short for Adobe's 'Portable Document Format,' comprises a large number of different versions of the format. This diversity of formats is further increased by innumerable third-party products which emulate the Adobe Acrobat Writer and produce PDF files of good or lesser quality.

Unfortunately, version 6.0 of the Acrobat Reader is not as lean an application as earlier versions were. And it is true that sooner or later, i.e. when the latest version of the Acrobat Writer has established itself, you will have to use the most recent version of the Acrobat Reader. However, the principle we want to draw attention to here has remained more or less the same from the Acrobat Reader version 4.0 onwards. An older version may possibly serve our purpose better.

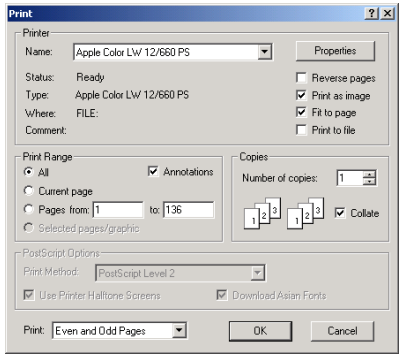
38.1.3.1 Difficult PDF files

PDF files often create difficulties. In the Acrobat Reader (but also in other products) there is an option to print these PDF files directly as image files. In the following an example: Open any PDF file you like (or the Archivista manual as in the example below).



Go to the menu 'Print...' (do not use the printing icon in the toolbar) and have a look at the printing options. For our purposes it is essential that a) you use the Archivista printer and b) you tick the option 'Print as image.' With the latter option each page is rasterized as image in the first place (no attempt at recognising fonts is made, for example). This way rasterising takes a bit longer but the fonts are interpreted more clearly than when we rasterize without this option.

Depending on the Acrobat Reader version you use you may have to search for this option in different places, as not all Acrobat versions are the same. If you do not find it straight away, please consult the Acrobat Reader manual.



Now print the file to the Archivista printer as usual and you will get images of high quality, especially with color publications.

38.1.4 Working with web pages

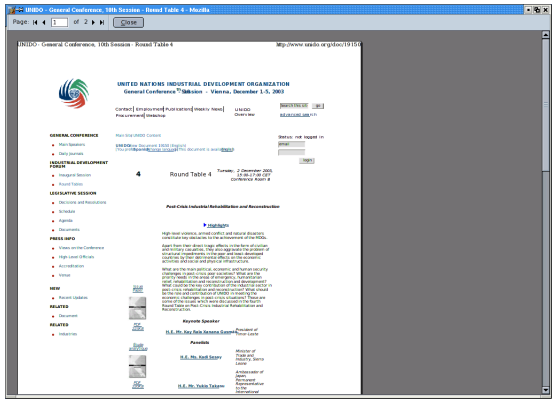
Archiving web pages is often a troublesome task. The reasons are manifold and we cannot go into all of them. We can discuss only a few major obstacles and how to remove them.

First of all we should not forget that the major hindrance to simply saving a web page as file onto our hard disk is that web pages normally consist of several files. Sometimes web pages are put together dynamically as when its contents are retrieved from a database.

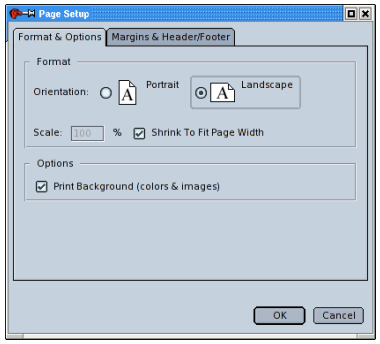
The easiest solution is if the web page we want to archive possesses a link to a print version. Some web pages offer this possibility because printing is explicitly desired.

If there is no print version available we should make sure that in terms of browsers there is not only the Internet Explorer installed but also Mozilla or Netscape (both rely on the same technology). Normally we get better results with Mozilla but in some cases Microsoft's Internet Explorer yields a more satisfying print output. If both fail, try a third type of browser (like Opera): it might give you the result you want.

Let us take a look at a web page in the print preview. It is not what we expect because the background is turned off and the preview's orientation is 'portrait.'

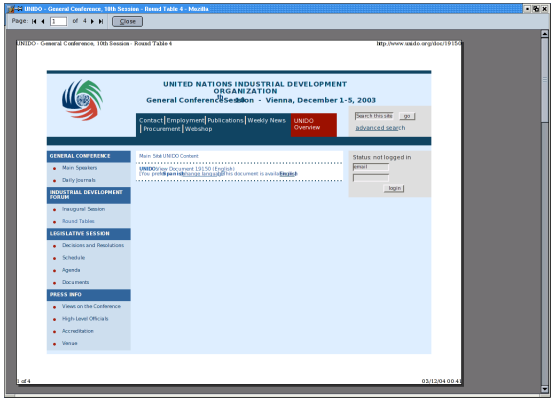


In 'Page Setup' we can change these values:

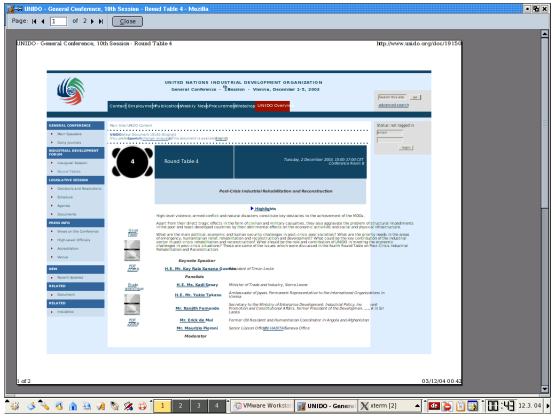


You find these options either directly in the 'Print preview' or, depending on the version of the browser you are using, in menu 'File.'

After these adjustments the page will look much better although it may happen that the contents have been miraculously moved to page two.



We can either start printing from page two or change the zoom factor to get a better result.



As soon as the print preview looks acceptable we print the page with the Archivista printer. The subsequent import of the page should not create any problems.

Part XI

Technical Details

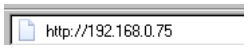
39 External SQL calls (WebClient)

For automated access to a particular database via WebClient we advise to follow the procedure and the conventions below.

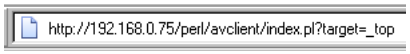
In principle, access via WebClient (e.g. to an ArchivistaBox) is defined by its IP address (or DNS name resolution). To access a specific database with a specific user, for example, it is possible to simply add some information to the IP address.

39.1 External login

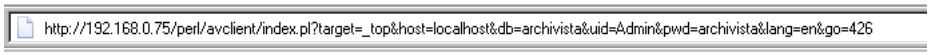
When the address of an ArchivistaBox is entered in the browser,



and 'Enter' is pressed, the IP address will automatically be complemented with a few internal parameters.



To this string of characters some other parameters can now be added in order to log in directly to a certain depth of the archive. The following, complete entry expresses the way with which the user reaches the view of the archive he needs or wants without him/her having to go through the individual steps manually. She/he wants 'localhost' for host, 'archivista' for database (with user 'Admin' who possesses password 'archivista'), and she/he wants an English layout but only document 426 to appear in the list.



The conventions are:

host	host computer
db	database
uid	user name
pwd	password
lang	language; the entry is en for English and de for German
go	go directly to document; here a document number is required

The order of the expressions between the & signs does not play a role. You can, for example, enter the password before the host. However, attention must be paid to making all necessary specifications: host and database are compulsory. 'lang' for language governs the layout, which means that for the sake of display it is compulsory, too. 'go' is not compulsory.

39.2 Extended access

The ArchivistaBox 2008/III (and higher releases) can be used to control the entire WebClient externally as well as carry out registration and file selection automatically. This means that it is now possible to control the WebClient 'remotely' via another (web) application. Firstly, this application has to connect to the WebClient (see 39.1). The application can then call all the functions of the WebClient that are available in 'manual' operation.

All editing functions are available here. Files can be deleted, scanning is possible and indexing can be automated. In order to gain a better understanding of external access, it is important that you are aware of the various forms (main/page view) and views (display, search, edit) since not all of the commands are available in all views.

The WebClient has two forms (views): the main view (`go_main`)and the page view (`go_page`). There are three different modes in these two forms (statuses): display (`go_view`), search (`go_search`) and edit (`go_edit`) documents. Before the individual commands are described, these different form views are shown here in brief form:

39.2.1 Forms

39.2.1.1 Main view

You can use `go_main` to switch to the main view at any time. Here, the system displays a table with the data records that are currently activated in the overview as well as the active data record in the detail view and a thumbnail.

39.2.1.2 Page view

You can use `go_page` to activate the page view. This allows you to display pages in enlarged form. You are also provided with functions for rotating pages and carrying out post processing activities.

39.2.1.3 Switching between the page view and the main view

You can use `go_pageswitch` to switch between the main view and the page view.

39.2.1.4 Switching to photo mode

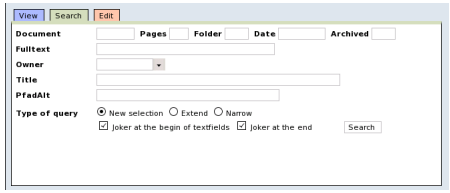
In photo mode, the pages are displayed as small thumbnails. Thus, the command for this mode is `go_thumbs`.

39.2.2 Modes (statuses)

39.2.2.1 Display

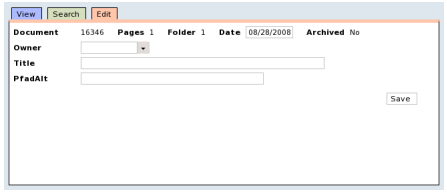
You can use `go_view` to switch to display mode. This displays the values that are currently saved in the database.

39.2.2.2 Search



You use `go_search` to activate the search form. You use `go_query` to trigger a search process and the fields are to be displayed with `fld_Field1=Value1&fld_Field2=Value2`.

39.2.2.3 Edit



You use `go_edit` to access editing mode. You can use `go_update` to edit the field values of actual files by entering the fields with `fld_Field1=Value1&fld_Field2=Value2`. However, this is only possible if editing rights exist in the respective file.

39.2.3 Logging in and logging out

39.2.3.1 Login

The login procedure is described earlier in the handbook under 39.1. Once you have logged in, the parameters are not required until the end of the session (i.e. when you logout).

39.2.3.2 Logout

You use `go_logout` to end the current session. The system then displays the form for logging in.

39.2.4 Navigation within files

39.2.4.1 Select all data records

You use `go_all` to display all files for the active session. This is also the case when the session is started and the first files are displayed.

39.2.4.2 Display previous data records

You can use `go_docs_prev` to display the previous data records in the selection. This is only useful if the first file is not already activated.

39.2.4.3 Previous data record

You use `go_doc_prev` to select the previous data record.

39.2.4.4 Selecting a data record

You can use `go_select` and `selectnr=xx&selectpage=yy` to activate a specific file and page from the **current selection**. You have the option of specifying a page but if you do not, the first page is always activated. This function is used primarily if you want to display a specific document or page after a search query (`go_query`).
The following example describes the selection of page 50 from document 4 (0 is the first document) from the **current selection**.

```
http://localhost/perl/avclient/index.pl?
go_select&selectnr=3&selectpage=50
```

➲ The above example uses the address 'localhost'. This only applies if you are working locally on an ArchivistaBox. In all other cases, 'localhost' must be replaced with the IP address or the name of the computer (domain).

39.2.4.5 Next data record

You use `go_doc_next` to activate the next file.

39.2.4.6 Display next data records

You use `go_docs_next` to display the subsequent files within the selection.

39.2.5 Navigation within pages

39.2.5.1 Display first page in document

You use `go_page_first` to select the first page in the current document.

39.2.5.2 Previous page

You use `go_page_prev` to select the previous page in the current document.

39.2.5.3 Jump to a specific page

You use `go_page_jump` to activate a specific page directly.

39.2.5.4 Subsequent page

You use `go_page_next`to select the subsequent page in the current document.

39.2.5.5 Display last page

You use `go_page_last` to select the last page in the current document.

39.2.6 Search in archives

39.2.6.1 Trigger search queries

You use `go_query` to trigger the search process.

To do so, you have to specify a field with `fld_Fieldname=Value` (or multiple fields) and/or a full text query with `fulltext=value`.

You use `searchmode=xxx` to determine the type of query. You have the following options for this:

- `searchmode=new` for a completely new query
- `searchmode=or` for extending the search query (or)
- `searchmode=and` for extending the search query (and)

Throughout the entire query, you can specify whether you want the system to use wildcards for the search or not. If you use `jokerstart=1`, a wildcard is used at the start of the field and if you use `jokerend=1`, a wildcard is used at the end of the field. The required values merely have to be somewhere within the field you require for both of these options.

The following example illustrates the search process:

```
http://localhost/perl/avclient/index.pl?
go_query&searchmode=new&
fld_title=Meier&fulltext=vacuumcleaner&
jokerstart=1&jokerend=1
```

When you start a new search query, this query searches the title field for 'Meier' with the additional prerequisite that the word 'vacuumcleaner' also has to occur in the full text. Wildcards are used at the start and end of fields meaning that all 'Meierhans' and 'Hansmeier' entries are also found.

39.2.6.2 Sorting search queries

You can use the `go_order_desc` and `go_order_asc` commands to sort a query in descending or ascending order for a specific field. To do so, you also have to specify the desired field accordingly with `orderfield=Fieldname`.

For example:

```
http://localhost/perl/avclient/index.pl?
go_order_asc&orderfield=Date
```

When you use this command, the current selection is sorted in accordance with the 'Date' field.

39.2.7 Editing files

By using go_update and entering fields (e.g. fld_ Title =Meier), it is possible to carry out post processing for the files. It is only possible to edit fields if the corresponding rights exist in the required file. If this is the case, editing is carried out without any further queries from the system.
For example:

```
http://localhost/perl/avclient/index.pl?
go_update&fld_Title=Meier&fld_lNo=1001
```

This command sets the 'Title' field to 'Meier' and '1001' is entered in the 'No' field.

39.2.8 Manipulating the display

39.2.8.1 Rotate left

You use go_rotate_left to rotate the current page to the left.

39.2.8.2 Rotate by 180 degrees

You use go_rotate_180 to rotate the current page by 180 degrees, that is, to turn it upside down.

39.2.8.3 Rotate right

You use go_rotate_right to rotate the current page to the right.

39.2.8.4 Zooming in and out on pages

➲ The following functions are only available in the page view since you can only zoom in or out on a page there.

You can use go_zoom&zoom=x to change the display continuously. Values between 0.1 and 1 (10 to 100 percent) are possible for x.
You use go_zoom&zoom=0 to display the page in accordance with the available screen area.
You can also use go_zoom_in and go_zoom_out to zoom in or out on the document depending on the current view.

39.2.9 Further commands (action)

Under normal circumstances, some commands can only be accessed in the WebClient if you activate editing for a page or file beforehand. You can use go_action&action=xxxto access these commands, regardless of whether you activate editing beforehand. The only condition is that the file can be changed, i.e. the relevant user is authorised to do so.

39.2.9.1 Deleting files

You can use go_action&action=delete to delete the current file.
You can delete multiple files directly by adding seldocs=x,y,z and using the relevant file numbers in place of x, y and z. You can delete a maximum of 16 files directly.
For example:

```
http://localhost/perl/avclient/index.pl?
go_action&action=delete&seldocs=1,4
```

Files 1 and 4 are deleted in the active database.
➲ Deletion is carried out without any further queries from the system. However, files to be deleted must not be blocked (undergoing processing) or already archived. Files such as these cannot be deleted.

39.2.9.2 Publishing files

You can use `go_action&action=publish&owner=xxx` to publish archives.

➲ For?owner', you have to enter the desired owner or the relevant group. If you want to publish the file for all users, you can do so by using `owner=[ALL]`. You can also use `seldocs=x,y,z` to publish multiple files directly.

39.2.9.3 Undoing published archives

Publishing can also be undone. You can use `go_action&action=unpublish` to do this. Once again, you can use `seldocs=x,y,z` to process multiple files simultaneously.

39.2.9.4 Combining files

You can use `go_action&action=combine` to combine two or more files (i.e. they are merged to form one file). You must enter `seldocs=x,y` here so that two files can be transferred to one file. The target file in this case is the active file.

39.2.9.5 Scanning files

You can use `go_scan` to scan a new file. A prerequisite for this is that a scanner is connected to the respective ArchivistaBox.

In addition, you can use `append_to=xxx` to carry out a scan in an existing file. You enter the desired file number in place of 'xxx'.

You use `scandef=y` to specify the desired scan definition.

Furthermore, you can also include field information with a file during scanning. You can do this by using `meta=Fieldname:Value`. Here, multiple fields can be addressed with multiple `meta` blocks.

The following example is intended to clarify the various possibilities during scanning:

```
http://localhost/perl/avclient/index.pl?
go_scan&append_to=44&scandef=3&
meta=Titel:Meier&meta=OrderNo:1001
```

If the above command is submitted as a whole line under 'localhost', the scanning process is triggered and the scanned pages are added to file 44. The third scan definition is used in this instance and the file then receives the value 'Meier' in the 'Title' (you have to user the german name 'Titel') field and the value '1001' in the 'OrderNo' field.

39.2.9.6 Delete page

You can use `go_action&action=deletepage` to delete a single page.

39.2.9.7 Saving changes to a page

You can use `go_action&action=savepage` to save a page that has been, for example, rotated.

39.2.9.8 Changing the OCR definition for a page

You can use `go_action&action=ocrpage!x` to assign a new OCR definition to a page. This ensures that this page is processed with the newly assigned OCR definition the next time text recognition (OCR) is called.

39.2.9.9 Setting OCR to complete for a page

You use `go_action&action=ocrdone` to set a page to complete for the OCR recognition. In other words, it is not used for the text recognition.

39.2.9.10 Excluding a page from OCR

You use `go_action&action=ocrexclude` to exclude the current page from any text recognition.

➲ You should use this option if certain pages within files only contain images since the OCR text recognition does not provide usable text and does not work quickly in the case of images.

39.2.9.11 Transferring the OCR definition to the file

You use `go_action&action=ocrselectdoc` to transfer the OCR setting of the current page to the entire file.

39.2.9.12 Transferring the OCR definition to the selection

Yo use `go_action&action=ocrselectdocall` to transfer the OCR definition of the current page to all of the files displayed.

39.2.10 Printing pages

You can use `go_print&print_from=x&print_to=y` to print a file. However, this is only possible if a postscript printer is installed on the respective ArchivistaBox and is activated for the users in Archivista WebAdmin.

39.2.11 Output format

Normally, when the files are accessed externally, the normal view of the WebClient is displayed. In order to make it as simple as possible to call information, a few output formats are provided so that specific information can be obtained directly. These display modes are outlined in the following.

39.2.11.1 Obtain page text

You can use `go_text` to obtain the current page of a file as a text file.

39.2.11.2 Obtain PDF file

You can use `go_pdf` to obtain the current file as a PDF file.

➲ Note: depending on the size of the file, it may take some time to obtain a file as a PDF file since the whole PDF file has to be created and delivered first.

39.2.11.3 Display an image file

You can use `go_image_doc_page_width_height` to call a page directly as an image. Here, `doc` represents the file number, `page` represents the page number, `width` represents the maximum width and `height` represents the maximum height.
For example:

```
http://localhost/perl/avclient/index.pl?
go_image_2108_2_1000_1000
```

Page 2 of file 2108 is displayed. The image is displayed with a maximum height and width of 1000 pixels.

39.2.11.4 Display Ajax list

You can use `go_ajax&ajaxfield=Fieldname&ajaxval=xx` if you want the system to read the values from Ajax fields (these are fields in the WebClient for which a window with corresponding values is opened when you perform a search or carry out editing).
For example:

```
http://localhost/perl/avclient/index.pl?
go_ajax&ajaxfield=Owner&ajaxval=a
```

In this example, all available owners for a file are returned. This information could, for example, be displayed as follows:

```
<ul><li>Admin</li>
<li>Admin1</li>
```

39.2.12 Return values as an HTML form

Normally, when the files are accessed externally, the normal view of the WebClient is displayed. Since this display contains extremely complex forms, you can use `frm_Fieldname=1` to ensure that only specific fields are shown in a simplified display.
For example:

```
http://localhost/perl/avclient/index.pl?
go_query&searchmode=new&fld_pages=4&
frm_filenumber=1&frm_pages=1
```

All files that contain precisely four pages are to be called. The 'Document number' and 'Pages' fields are to be displayed. Unsurprisingly, 'Pages' always contains the value '4' in the form:

```
<input type="hidden" name=" filenumber _1" value="1251">
<input type="hidden" name=" pages _1" value="4">
<input type="hidden" name=" filenumber _2" value="1229">
<input type="hidden" name=" pages _2" value="4">
<input type="hidden" name=" filenumber _3" value="1228">
<input type="hidden" name=" pages _3" value="4">
...
```

39.2.13 Final example

To round off, the following program example illustrates a simple session with the WebClient in the Perl language:

```
use strict;
use LWP::UserAgent;
use HTTP::Cookies;
my $server = "http://localhost/perl/avclient/index.pl";
my $connect = "?host=localhost&db=archivista&uid=Admin&pwd=archivista";
my $www = LWP::UserAgent->new; # new www session
$www->cookie_jar(HTTP::Cookies->new('file'=>'/tmp/cookies.lwp','autosave'=>1));
my $res = $www->get("$server$connect"); # connect to webclient
if ($res->is_success) {
  if ($res->content) { # we got login
    my $cmd = "?go_query&fld_Seiten=3"; # search for 3 pages
    $res = $www->get("$server$cmd");
    my $cmd = "?go_doc_next";
    for (my $c=1;$c<=18;$c++) { # navigate 18 times to next doc
      $res = $www->get("$server$cmd");
    }
  }
}
```

The above example creates a connection with the WebClient, executes a search query in which all files with three pages are activated and then navigates through 18 files. Furthermore, when a script is called externally, it is important that the 'cookies' are saved. This is the only way in which further 'commands' can be transferred to the WebClient.

39.2.14 Transferring files to the archive

Since it is possible in the WebClient to transfer a file (PDF or graphic file) to the archive server, it is also possible to do this via external control. In principle, the `go_action&action=upload` command can be used for this. However, when you upload data, it is necessary to switch to 'POST' mode during communication with the WebClient (this is different from the GET mode that is normally used).
During the upload, you can include the `uploadocr=x` and `uploadbits=y` options in order to determine OCR definitions as well as scanning. Alternatively, you can also determine a scan definition containing the processing parameters directly during the upload. You use `uploaddef=x` to select a specific scan definition, although the first scan definition is activated with `0`.
The following example demonstrates uploading a file to the web server in the Perl programming language. The comments should be of assistance even if you carry out the upload in a different programming language.

```
#!/usr/bin/perl

# wcupload.pl -> demo script for uploading a document via web client
# (c) 2008 by Archivista GmbH, Urs Pfister
```

```
use strict;
use LWP::UserAgent; # we work with UserAgent (our batch web browser)
use HTTP::Cookies; # we need to work with cookies
use HTTP::Request::Common qw(POST); # the post method must be imported

my $fin = "/home/archivista/documentation_de.pdf"; # demo doc (our manual)
my $pages = "1-5"; # pages to extract (so we don't wait too long)
my $fout = "/tmp/eins.pdf"; # the finally file we want to import
if (!-e $fout) { # create file with pdftk if it does not already exist
  system("pdftk $fin cat $pages output $fout");
}

# server we use (link to webclient)
my $server = "http://localhost/perl/avclient/index.pl";
# connection string (host,db,user,password)
my $connect = "?host=localhost&db=archivista&uid=Admin&pwd=archivista";

my $www = LWP::UserAgent->new; # new www session
# save the cookie for the corrent session
$www->cookie_jar(HTTP::Cookies->new('file'=>'/tmp/cookies.lwp','autosave'=>1));
my $res = $www->get("$server$connect"); # connect to webclient
if ($res->is_success) {
  if ($res->content) { # we got login
    # now upload a file, we use request method with POST
    my $res = $www->request(POST "$server",
      Content_Type => 'form-data', # multipart/form-data
      Content => [ # structure for our file
        MAX_FILE_SIZE => 134217728, # max. size (WebClient won't accept more)
        upload => [ $fout, $fout ], # file to upload, file name to use
        go => 'go_action', # we need to call the go_action command
        action => 'upload', # inside of go_action we need to use upload
        uploadbits => 1, # 1=black/white, 8=gray, 24=color
        uploadocr => 27, # the desired ocr def (1-x, 27 does mean: no ocr)
        meta => "Titel:450", # filling in some meta keys is no problem
      ]
    );
    if ($res->is_success) { # if we got a succes, file is uploaded
      print "file $fout uploaded\n";
    }
  }
}
```

➲ Note: The above example runs in every ArchivistaBox. In principle, however, the WebClient can be addressed by any computer and it is even possible to work with encrypted access (HTTPS). You can find the program on the ArchivistaBox CD under `/home/cvs/archivista/jobs` under the name `wcupload.pl`.

39.2.15 Working with archived mail

Archived mail files can also be copied back to the original mail server from the WebClient. To do so, you use the `go_mail_xxx` command, where 'xxx' represents the file number. You can use `go_zip_xxx` to download the mail message as a zipped file.

39.2.16 Working with Office files

Archived Office files can also be restored from the WebClient. To do so, you use the `go_file_xxx`, command, where 'xxx' represents the file number. You can use `go_zip_xxx` to download the Office file as a zipped file.

39.2.17 Determining the WebClient version

Each page that is delivered by the WebClient contains a hidden 'avversion' form variable that contains the current date of the release (e.g. 20080222). This allows the calling application to test whether a sufficiently updated version is available.

40 External SQL calls (RichClient)

Archivista Business and Enterprise provide an interface for accessing individual documents within Archivista directly from third party applications. To do this you need the 'avstart.exe' auxiliary program. As soon as the program is started, a database query is carried out in Archivista. Any documents found are displayed directly in Archivista ('avstart.exe' is terminated).

➲ Note: The auxiliary program 'avstart.exe' is capable of starting Archivista automatically if Archivista is not up and running already. For this purpose you must create the file 'avstart.dat.' The first line of this file contains the program name of Archivista (Av5.exe). The second line may contain the directory path (e.g. 'c:\Program Files\Av5e) or may be empty. The third line must consist of the start-up parameters of Archivista, which look approximately like the following: '* s localhost;archiv;SYSOP.' Please note in particular that you must put an asterisk followed by a space followed by a (lower case!) 's' followed by a space and then enter host, database and user separated by semicolons. (The user can also be omitted.)

The following control parameters may be used when accessing 'avstart.exe': SQL (SQLEQUAL), SQLLIKE, FIRST, PREVIOUS, NEXT and LAST (capital letters required). The commands SQL and SQLLIKE are special in that when accessing 'avstart.exe' and SQL or SQLLIKE, additional field names and values are also required. Here are some examples:

c:\Program Files\avstart.exe SQL;Titel,Meier

c:\Program Files\avstart.exe SQLEQUAL;Titel+,Meier

In both cases, a query will be run in the 'Title' field, i.e. the documents activated will be those which have exactly the word 'Meier' in the 'Title' field.

c:\Program Files\avstart.exe SQLLIKE;Datum-;Titel,Meier

In this case, the files activated will be all those in which 'Meier' occurs, in any position, i.e. 'Hansmeier' or 'Meierhans' will also be located. In this example the hits will be sorted by date, in descending order (because of the minus sign in the field name), and then according to the title field.

Note: Up to five fields can be used for selection/sorting. Each unit is separated by a semicolon (;). The field name and value are separated with a comma (,). If no field value is assigned, the field will be used only for sorting. If the symbols '\', ';' and ',' are to be used, they must be specified as '\-', '\,' bzw. '\.'

➲ **Note regarding activation:** Installing Archivista 2009/I does not automatically activate the program 'avstart.exe'. The relevant program file is initially installed in the program directory as 'avstart.eee'. Only when you rename the file 'avstart.eee' as 'avstart.exe' will you be able to work with this functionality. What is more, you need a licensing file. On acquiring this module you receive the licensing file on a CD or via e-mail and you must copy it into the Archivista program folder before you can make full use of 'External SQL calls.'

41 Database structures

41.1 The most important Archivista tables

In the following we would like to introduce you to the most important tables of an Archivista database. Please note that table and field names are in German as yet. For a conversion table of field names please see chapter 35.6.

- **Table 'archiv':** Each record corresponds to an Archivista document. The field 'Laufnummer' (English: 'counter') administrates the document number. The field 'Gesperrt' (English: 'locked') is essential in that each time a user tries to change a document the system checks this field to find out whether or not somebody else is already working on the same document. If the field 'Gesperrt' is empty, the user's ID is entered in the field and he is free to make changes. After he has finished the content of the field 'Gesperrt' is erased again.
- **Table 'archivseiten':** Each record corresponds to a page in a document. The necessary link between documents and pages is saved in the field 'Seite' (English: 'page'): the last three digits denominate the pages of a document while the preceding digits stand for the document number. Let us make an example: '2412004' can be parsed into document number '2412' and page number '4.' It is important to know that there is no relational mapping between the two tables as yet. That is, an application must take care, for example, that all pages in table 'archivseiten' are deleted, the page counter in table 'archiv' (field 'Seiten') is updated and the document itself is deleted (if there are no longer any records in table 'archivseiten').
- **Table 'parameter':** In this table you find all settings which play a role in any Archivista database. Most relevant are certainly the entries regarding the archive directories. These entries begin with 'Pfad ...' (English: path).
- **Table 'user':** In this table you find the user specific settings. For more information see 41.3.

41.2 Archiving concept

41.2.1 Tables and Fields

41.2.1.1 Table 'archive'

In table 'archive' the individual documents are stored. The keywords, the number of pages and the folder number are to be found here, for example, but not the image files or texts of the individual pages.

```
mysql>
mysql>
mysql> select Akte,Seiten,ArchivArt,Erfasst,Archiviert,BildInput,BildIntern,QuelleIntern,BildInputExt,BildAExt,QuelleExt from archiv or
der by Akte desc limit 3;
+------+--------+-----------+---------+------------+-----------+-----------+--------------+--------------+----------+-----------+
| Akte | Seiten | ArchivArt | Erfasst | Archiviert | BildInput | BildIntern | QuelleIntern | BildInputExt | BildAExt | QuelleExt |
+------+--------+-----------+---------+------------+-----------+-----------+--------------+--------------+----------+-----------+
|   77 |    230 |         1 |       1 |          0 |         0 |         0 |            0 |              |          |           |
|   76 |     16 |         1 |       1 |          0 |         0 |         0 |            0 |              |          |           |
|   75 |      1 |         1 |       1 |          0 |         0 |         0 |            0 |              |          |           |
+------+--------+-----------+---------+------------+-----------+-----------+--------------+--------------+----------+-----------+
3 rows in set (0.00 sec)

mysql>
```

Let us take a look at the relevant fields of table 'archive':

- Dokument: unique number for document (1-x)
- Seiten: Number of pages per document (0=none,1-640 pages)
- ArchivArt: 0=BMP/ZIP (obsolete), 1=TIF, 2=PNG, 3=JPG
- Erfasst: 0=no pages added, 1=document contains pages

- Archiviert: 0=document is not archived, 1=document is archived
- BildInput: 0=image in 'input' folder, 1=image in table 'archivbilder'
- BildIntern 0=no web preview images, 1=web preview images
- QuelleIntern: 0=no source file, 1=source file
- BildInputExt: file extension in capital letters (TIF,PNG,JPG)
- BildAExt: file extension for web preview images (TIF,PNG,JPG)
- QuelleExt: file extension for source format archiving (e.g. DOC,PDF)

41.2.1.2 Table 'archivseiten'

Table 'archivseiten' contains the information regarding the text recognition of a page, i.e. we find the extracted text as well as the OCR definition belonging to the page. Note that the archiving process leaves table 'archivseiten' untouched.

```
mysql> describe archivseiten;
+--------------+-------------+------+-----+---------+-------+
| Field        | Type        | Null | Key | Default | Extra |
+--------------+-------------+------+-----+---------+-------+
| Seite        | int(11)     |      | PRI | 0       |       |
| Ausschliessen | tinyint(4)  |      | MUL | 0       |       |
| Erfasst      | tinyint(4)  |      | MUL | 0       |       |
| Schlüssel    | varchar(56) |      |     |         |       |
| Text         | mediumtext  | YES  | MUL | NULL    |       |
| Zipped       | tinyint(4)  |      |     | 0       |       |
| Indexiert    | tinyint(4)  |      | MUL | 0       |       |
| OCR          | int(11)     |      |     | 0       |       |
| Notes        | blob        | YES  |     | NULL    |       |
| ScreenQuality | int(11)     |      |     | 0       |       |
+--------------+-------------+------+-----+---------+-------+
10 rows in set (0,06 sec)

mysql>
```

The following fields are relevant for us:

- Seite: reference to document and page of table 'archive' (document*1000+page)
- Ausschliessen: do not treat page with OCR
- Erfasst: page already treated with OCR
- Text: memo field for various information
- Indexiert: shows whether an index has been done for this document (obsolete)
- OCR: definition (0-x) making the desired language strings available
- ScreenQuality: reduction factor for screen copy: 0-50

41.2.1.3 Table 'archivbilder'

Table 'archivbilder' houses the individual image and source files.

```
mysql> describe archivbilder;
+-----------+------------+------+-----+---------+-------+
| Field     | Type       | Null | Key | Default | Extra |
+-----------+------------+------+-----+---------+-------+
| Seite     | int(11)    |      | PRI | 0       |       |
| Bild      | mediumblob | YES  |     | NULL    |       |
| BildA     | mediumblob | YES  |     | NULL    |       |
| BildInput | longblob   | YES  |     | NULL    |       |
| Quelle    | longblob   | YES  |     | NULL    |       |
| BildX     | int(11)    | YES  |     | NULL    |       |
| BildY     | int(11)    | YES  |     | NULL    |       |
| BildAX    | int(11)    | YES  |     | NULL    |       |
| BildAY    | int(11)    | YES  |     | NULL    |       |
| DatumA    | datetime   | YES  |     | NULL    |       |
+-----------+------------+------+-----+---------+-------+
10 rows in set (0,07 sec)

mysql>
mysql>
```

Let us take a look at the relevant fields of table 'archivbilder'

- Seite: reference to document and page of table 'archive' (document*1000+page)
- Bild: first image file existing in reduced web quality
- BildA: currently not used
- BildInput: original image file, i.e. page as captured

- Quelle: source file (either from first page or each page individually)
- BildX, BildY: pixel data of file in field 'Bild' (optional)
- BildAX, BildAY: pixel data of file in field 'BildInput' (optional)

41.2.2 Archiving process

The objective of the archiving process is to save the added documents in a way that they cannot be altered anymore and that they remain readable from a long-term perspective.
That image data become unalterable is effected by the field 'Archiviert' in table 'archive' showing '1': any manipulation of the image in an Archivista application is no longer accepted.
Long-term readability of the image data is achieved by the fact that in any case they are saved in a certain folder structure and that the images thus created, or the corresponding folders, respectively, are written to non-rewritable media (CDR/DVDR).

41.2.2.1 Purpose of folder number

The field 'Folder' of table 'archive' serves to store the information in which folder a specific file (image file) can be found. The folder number is managed by the system. The last, i.e. most recent folder number is managed in table 'parameter'. There the field 'Name' has an entry 'ArchivOrdner'. The number found there is the currently used folder number.
In principle each document is allocated the current folder number when the document is being created. The folder number refers to the corresponding archiving folder. Example: if document 84 shows folder number 10, it will be archived in folder 'ARCH0010' if the previous documents have found room too.
Important for the archiving process are the following additional entries in table 'parameter':

- ArchivAutomatisch: 0=manual (obsolete), 1=automated archiving process
- ArchivMByte: 100-10000 (value=number of MByte per archive folder)
- ArchivDateien: 100-10000 (value=number of files per folder)

In 'ArchivMByte' one can determine how many MByte are to be saved in a folder (e.g. 600 MByte), in 'ArchivDateien' one can specify the maximum number of files (e.g. 2000 files) that are allowed in a folder.
In the following we give two examples.
First example: For the moment we do not know if document 84 can still be archived in folder 10 (name: ARCH0010). Currently, folder '10' is made up of 25 MBytes and 1789 files. The field 'ArchivDateien' shows 1800 files, our document 84 possesses 14 pages. Together with document 84 folder 'ARCH0010' would contain 1803 files. That is the reason why a new archiving folder with the name 'ARCH0011' is created and the current folder number goes up to 11.
Second example: Up to and including document 83 there are 1500 files in folder 'ARCH0010'. The folder consists of 589 MBytes, document 84 contains 220 files and is made up of 44 MBytes. Also in this case document 84 would be finally saved in archive folder 'ARCH0011'.

41.2.2.2 Calculation of file name

If a page is saved in the database itself the image file is not named. The file name is calculated on the basis of the field 'Document' and the corresponding page. It works like this:
The image file gets an 'A' in first position. The second to the sixth position are reserved for the document number, the seventh and eighth position gives the current page.

The following example will make this clear.

Document number 84: modulo 26 = 6 (G) remainder, 3 (D) results in: 000DG

Page number 12: modulo 26 = 12 (M) remainder, 0 results in 0M

Thus, the full file name (without extension) looks like A000DG0M.XXX. The extension of the file name is calculated on the basis of field ArchivArt:

- 0 = BMP/ZIP
- 1 = TIF
- 2 = PNG
- 3 = JPG

If field 'ArchivArt' displays the value '3' all this results in the file name A000DG0M.JPG.

41.2.3 Differences in the archiving process

41.2.3.1 Working with internal images (ArchivistaBox)

- The image files are immediately stored in table 'archivbilder' and field 'BildInput'.
- The source files are immediately stored in table 'archivbilder' and field 'Quelle'.
- The field 'BildInput' in table 'archive' shows '1'.
- During the archiving process the files are always moved to folder 'output'. The image files can remain in the database but they need not remain there.

41.2.3.2 Working without internal images (Archivista 5.x)

- When pages are added the image files are saved in folder 'input'
- Newly added source files are also filed in folder 'input'
- The field 'BildInput' of table 'archive' shows '0'.
- During the archiving process the files are moved from folder 'input' to folder 'output'.

41.2.4 Archivista folders

When Archivista archives documents it counts the number of pages present in the current directory. As soon as a certain total of pages is reached a new directory is automatically created and the surplus pages are copied there. The reason for this is that an operating system can only handle a certain number of files per directory efficiently. A new directory is created, too, when the size of the currently used directory exceeds the predetermined size of the external device. If no new directory were created in such a case, it could not be transferred later.

As mentioned previously the task of handling pages and preparing them for transferral to external media is taken over by Archivista. The only thing you should keep an eye on is the folder number. If after archiving the number in the field 'folder' is increased, a new folder has been created and the old one (only that one) may be transferred.

An example: The fields 'folder' of documents 1 to 34 show a 1 before archiving. After the archiving process the documents 29 to 34 have a 2 in their respective field 'folder'. Folder 1 is definitively full and may be transferred.

If there is an output directory under 'C:\Program Files\av416e\archive\output\' you will find the two subdi-

rectories ‘arch0001’ and ‘arch0002’ in them after the archiving process has been run. The first subdirectory, ‘arch0001’, can be transferred to an external storage device because it is complete (‘arch0002’ is incomplete and should be left alone).

➲ If you work with standard settings the folders are made to fit on CDR. You need not make adjustments to the parameters for archiving.

A further piece of advice: To be on the safe side with regard to data security you should make two copies of the complete folder that you want to transfer.

41.3 Clean login to Archivista databases

At the beginning is a clean login to the system. Basically you need the required MySQL rights to access an Archivista database. To allocate these rights you either create a user on the MySQL console or you use an ordinary Archivista user account.

The first alternative is probably easier to implement while the second gives you more flexibility. With the second alternative you can fall back on the existing user specific information, something you cannot do when using the first. The user specific information is to be found in table 'user.' This is the same for all Archivista databases.

When login is attempted through an Archivista client the system checks whether the user possesses access rights (reading rights as a minimum) for a particular Archivista database. The MySQL server consults its internal administration database 'mysql' and takes a look at the entries in table 'tables_priv.' If user ID and password are correct, the user may access the database. In a second step the system checks the user specific information in table 'user' of our Archivista database. Is there an entry which has the same information in fields 'user' and 'host' as the internal MySQL table 'tables_priv'? If this is the case, too, login may definitively take place.

41.4 File name of an external saved page

After we now know how to access an Archivista database we would like to show an exemplary case of how to locate bitmap pages with an SQL query and a few commands in Perl.

```
sub GetDateiName {
    my $modus = shift;  # A or B copy
    my $akte = shift;   # document number
    my $seite = shift;  # page in question
    return $modus . FileName($akte,5,) . FileName($seite,2);
}

sub FileName {
  my $eingabe = shift;
  my $anzahl = shift;
  my ($res,$res1,$out);
  while ($eingabe != 0) {
    $res = $eingabe % 26;
    $res1 = $res + 65;
    $out = chr($res1).$out;
    $eingabe = $eingabe - $res;
    $eingabe = $eingabe / 26;
  }
  $out = "0" x ($anzahl - length($out)) . $out;
  return $out;
}
```

The code above shows how the file name is calculated from the document number and the page number in that document.

To calculate the complete path of a page we in addition need the type of the document and the folder where it lies. The type of the document is to be found in field 'ArchivArt': 0=BMP/ZIP,1=TIF,2=PNG and 3=JPG. The folder number is to be found in field 'Ordner.'

One other thing must be paid attention to: has a document been archived or not? The field 'Archiviert' in table 'Archive' tells us. If yes, the page will no longer be found in the input folder but it will reside in the output folder. The folder of a document that has not yet been archived can be located by the following procedure:

a) it must be in the input directory b) its number is the result of taking the remainder of modulo 100 of a document number (e.g. the remainder of 1725 would be 25).

Two more examples:

Page in input folder: Document 1, page 1 with A copy and value '1' in field 'ArchivArt' you find in '..\input\in001\A0000B0B.TIF'.

Page in output folder: Document 2, page 3 with B copy and format 'JPEG' has already been archived and lies in folder '1' (field 'Ordner,' English: 'folder'). This page is to be found in '..\output\ARCH0001\B0000C0D.JPG'.

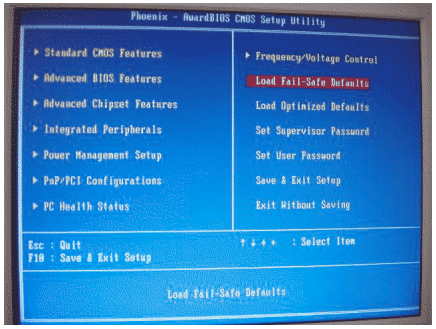
➲ The Archivista PublishingEdition offers more clues as to how to communicate with an Archivista database. It gives you the complete source code (including correct login) for communicating successfully with Archivista databases.

42 ArchivistaBoxes

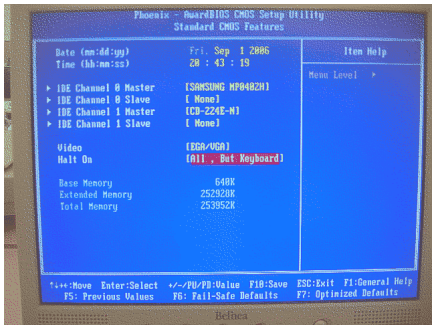
This chapter is only relevant for the technicians of Archivista Ltd.

42.1 BIOS settings

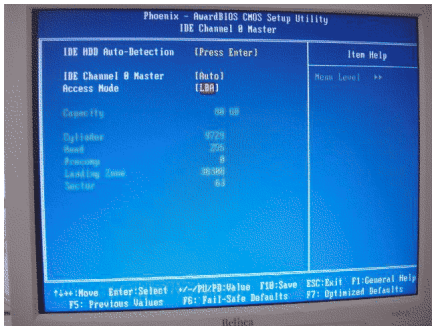
The following explanations refer to the default hardware settings recommended for our ArchivistaBox systems.



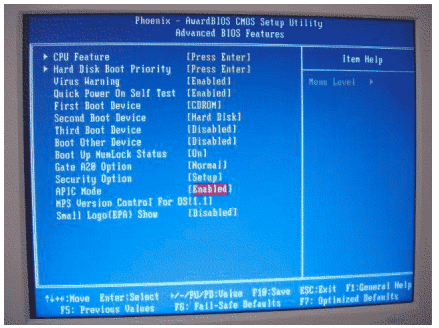
At first conservative BIOS settings are loaded.



Under 'Halt On' we make sure that one can work with the box even if there is no keyboard attached. Then, we move to the hard disk.

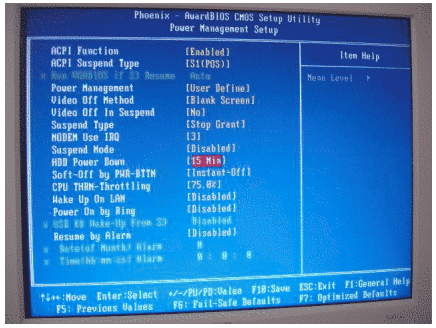


As access mode for the hard disk we choose in any case 'LBA' and ensure that the kernel can cope with disks beyond 8 GByte.

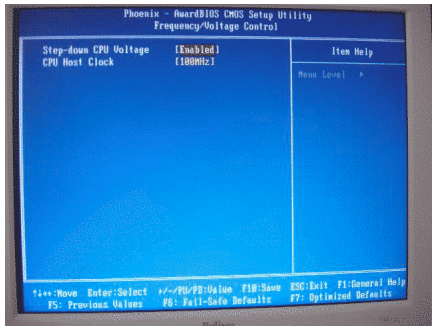


Finally, the APIC mode must be turned on. Otherwise the hardware (USB in particular) might not work properly.

To ensure that the ArchivistaBox does not 'freeze' we use RAM values as shown above.



Also the hard disk deserves a certain measure of protection.



Should it come to pass that the box is continuously under full load in high summer temperatures we reduce CPU voltage to avoid overheating.

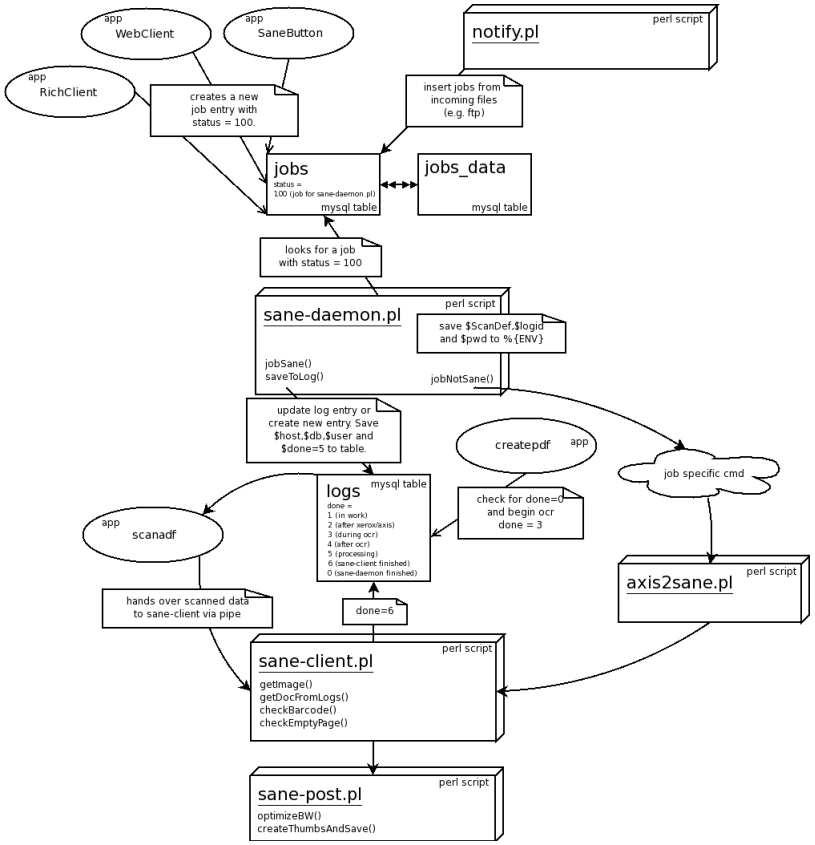
42.2 Hard disk partitions

The hard disk is configured in a way that automatically four partitions are created:

Inst1	hda1
Inst2	hda2
SWAP	hda3
Data	hda4

On the first two partitions (4 GByte each) a maximum of two different ArchivistaBox versions may be installed. The third partition is used as SWAP storage and on the fourth one the data are saved.

42.3 Processing of new jobs

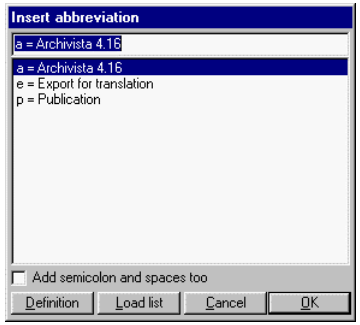


New jobs are inserted into the jobs table by an active application (e.g. WebClient) and flagged with status 100. Every two seconds the Sane deamon fetches all jobs with status 100 and processes them. It sets the job status to 120 and adds an adequate logs entry with an agreeing DONE value. Depending on the kind of

job it is, scanadf is run for a scanning job, or other applications are called up for other jobs. After having processed the job these applications hand over their data to Sane client, which sets the log entry to its final DONE value. Further, Sane client analyzes the prepared data and forwards them to Sane post.

43 RichClient

43.1 Keywording functionality



Practice has shown that many users seldom, if ever, use the 'F2' key ('Get abbreviation' in the 'Edit' menu). The abbreviation window (Ctrl+F2) is much more frequently used. However, since new abbreviations have to be created with 'F2' (or directly via the 'Insert abbreviations' table) many people have not been able to get into a habit of entering keywords at reasonable speed. For this reason, you can now create new abbreviations directly in the abbreviations window, under 'Definition'.

Also, when you are **selecting records,** the abbreviations window can be called up directly using the **'Abbreviations' button.**

At this point it is useful to recall that page text is intended mainly for full text queries, and keywords for selection. However sometimes the boundary between these two search strategies becomes blurred. If you want to copy information from the field 'Page text' to other fields, mark individual words in 'Page text' and then press 'F2' to transfer them to a particular, pre-defined text field. To specify the field, go to 'Parameters' and select 'General program parameters'.

43.2 Efficient entering of data

Before we describe certain procedures which simplify data entry we would like to mention the three standard shortcuts to cut, copy and paste bits of field content.

- **'Ctrl+Insert':** Copying of marked parts
- **'Shift+Insert':** Pasting of the parts previously copied or cut
- **'Shift+Delete':** Cutting of marked parts

In the following a few things which are good to know and which help you to work efficiently:

For often used words you best define an **'abbreviation'**, i.e. you define a code and when you want to enter that word you enter the code and press **'F2'.** More on this topic under menu 'Abbreviation' in chapter 31.2.

If you have to edit the same field in several records you should try the **'list'** or table of records (in the middle of the screen). You navigate more elegantly when working in the list and especially when you want to edit the field 'Marked' the list is useful. To edit a field you must go to that field and then press the Enter key before you can change anything. You can also change the width and height of the cells of the list to work more comfortably with it.

If you often add the same or similar records you should work with **'Marked parts on clipboard'** and **'Paste text from clipboard'**.

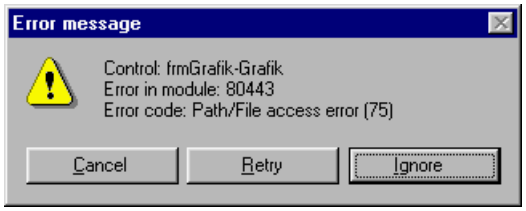
And last but not least you can enter the current date in **date fields** by pressing the space bar once.

43.3 Runtime errors

Archivista was developed with a view to high "crash resistance" and was tested extensively. Normally, no runtime errors should occur and you should be able to work without interruption.

If a runtime error occurs nevertheless, there are routines which prevent the program from crashing. The error is analyzed internally and a **general error message** appears.

In the following an example of a general error message.



Normally, you can pass over the error by clicking 'Repeat'. The command causing the error is run again. If again an error message occurs, click 'Ignore' and the command is passed over altogether.

If the error messages persist, click 'Cancel'. The current function is quit and the program asks you if you want to exit.

If the same error occurs again and again, please contact us in writing and describe what happens in detail. We also need the exact version number of your program (to be found in menu 'Help', menu item 'Info'), your customer number, a description of the hardware and software with which you work and the exact wording of the error message (all three lines).

➲ If error '3024' occurs, Archivista tries to load a non-existing database. In such a case delete file 'av5.dat' from your Archivista program directory.

Part XII

FAQs

44 FAQs regarding ArchivistaBox

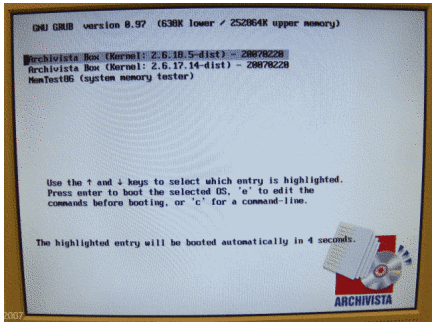
44.1 Is there a demo version of Archivista?

No, but you can install the GPL version at any time. If you want to use the commercial OCR option, we recommend that you take a look at the following association?s website: www.freearchives.ch.

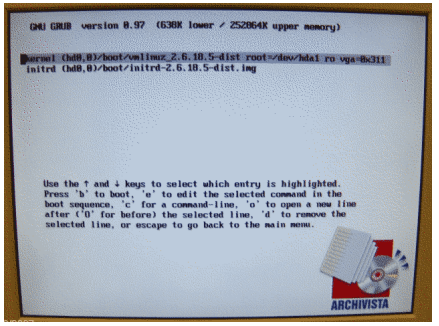
44.2 ArchivistaBox does not start up properly

After ArchivistaBox has been switched on normally the following message appears: 'System being started...'. Should it happen that the ArchivistaBox does not move beyond this screen you can adjust the start up screen temporarily so that you get more detailed information on the system.

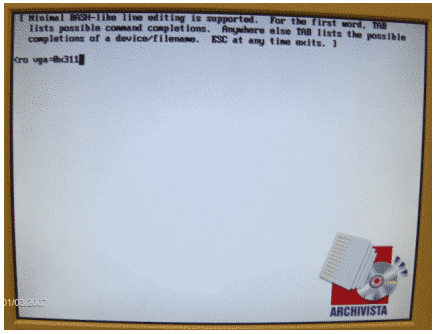
Shortly after turning on ArchivistaBox the following screen appears (may look slightly different). Now press 'e' to activate the highlighted (first) entry.


GNU GRUB version 0.97 (638K lower / 252064K upper memory)
Archivista Box (Kernel: 2.6.18.5-dist) - 20070228
Archivista Box (Kernel: 2.6.17.14-dist) - 20070228
MemTest86 (system memory tester)
Use the ↑ and ↓ keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
The highlighted entry will be booted automatically in 4 seconds.
ARCHIVISTA

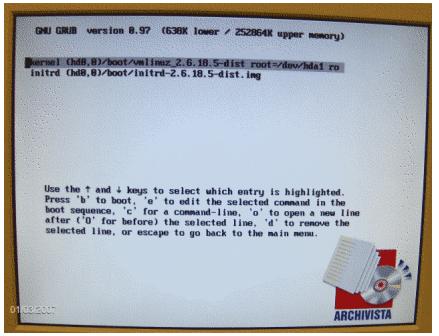
Press 'e' again to edit the entry.


GNU GRUB version 0.97 (638K lower / 252064K upper memory)
kernel (hd0,0)/boot/vmlinuz-2.6.18.5-dist root=/dev/hda1 ro vga=0x311
initrd (hd0,0)/boot/initrd-2.6.18.5-dist.img
Use the ↑ and ↓ keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
ARCHIVISTA

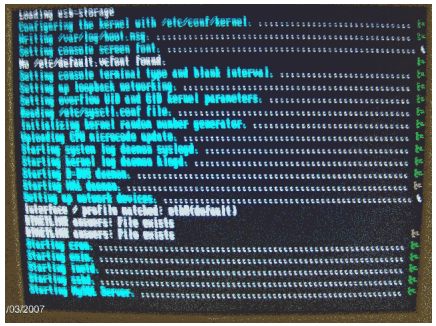
Delete the characters 'vga=0x311' at the very end of the line and press 'Enter'.


[Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits.]
ro vga=0x311
ARCHIVISTA

Start the booting process by pressing 'b'.


GNU GRUB version 0.97 (638K lower / 252064K upper memory)
kernel (hd0,0)/boot/vmlinuz-2.6.18.5-dist root=/dev/hda1 ro
initrd (hd0,0)/boot/initrd-2.6.18.5-dist.img
Use the ↑ and ↓ keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
ARCHIVISTA

Subsequent to this procedure you will not see the start up screen you normally see but line by line there will appear status information instructing you about system startup details. Where the start up process stops, there lies the problem: the last line will tell you where. If you have a support contract for ArchivistaBox, please take a picture of the screen by means of a digital camera and send the picture together with a description of the problem to the support department.



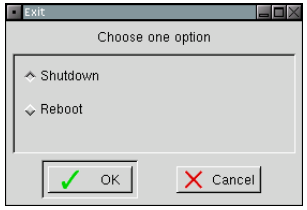
P.S. ArchivistaBox starts up fast. Please forgive us for not taking a sharper image.

44.3 Shutdown of ArchivistaBox

The ArchivistaBox is an archiving server designed for 7x24 hours' running. Nevertheless it may make sense under certain circumstances to turn it off.

Click the X in the upper right hand corner of the screen to change to the level of the ArchivistaBox settings. You will see a grey screen. Right-click your mouse and choose 'Exit' from the context menu that appears.

You are asked for the root password. Then, a new dialog box appears.



Click 'Shutdown' or 'Reboot' depending on your wishes.

The archiving server will shut down.

➲ Please note that there is no need to press any buttons for turning the ArchivistaBox off.

44.4 Installation/Update from CD

Tutorial on Installation (movie 2min) *4.2* **Tutorial on Installation (movie 2min)** *4.3*

It may happen that you want to carry out a system update. For this purpose you receive a CD with the new system. Follow the steps below:

- Close WebClient and WebAdmin
- Insert new CD
- Change to the ArchivistaBox settings by rightclicking your mouse. Choose menu 'Exit' and shutdown the ArchivistaBox as described in 44.3.
- Wait one minute
- Start ArchivistaBox by pressing the 'Power' button
- Now the ArchivistaBox will boot from the live CD. The startup process resembles the startup from hard disk but bear in mind that it is only the CD. The system is rather slow. After some time you are asked to define keyboard layout and

network configuration. For 'keyboard' please enter the layout that you need. But don't make an entry for 'network': skip this setting by hitting 'No'. After a moment of patience you will see the login screen of the WebClient. That is the end of the starting up process.

- Close WebClient
- Change to the ArchivistaBox settings by rightclicking your mouse. Now you should see the entry 'Install current system to a hard-disk partition' in menu 'System'.
- Run this command. See also 8.14.9. If you have reliable backups (a must in any case), you can choose the option `/dev/hda-Reformat_whole_disk`, i.e. you reformat the entire hard disk and are ready to copy the database and your data to the system later. If you do not want to touch the data existing on your hard disk, you choose a partition that has been already used (named 'Archivista' plus the date in brackets) and you overwrite simply the program or you use a pristine partition (that looks like `/dev/hda2-Not␣formated`) for the installation of the program.
- The installation process takes roughly 20 minutes. Please wait until the message appears that the installation is finished.
- The CD has done its work. Now shutdown the system. Please remember to use the default password 'archivista' for this.
- Start the ArchivistaBox by pressing the 'Power' button and open the CD drive as soon as you see the green lamp there. Remove the CD.
- Now the system boots from the newly installed hard disk.

44.5 Installation/update from USB-Stick

It is also possible to install the system by means of a USB stick. In this case proceed in the following way:

- Shut down the ArchivistaBox
- Attach USB stick, press 'Power' button
- Press the ESC key and keep it pressed until ...
- ... the system gets into a special mode where you can choose from which device the system should boot. Select the hard disk as 'Boot First Device'.
- A new screen appears where you can specify the 'Boot First Device' more exactly: choose USB-HDD.
- The rest of the procedure corresponds to what is described in 44.4 and 8.14.9, respectively.
- Remember to remove the USB stick only **after** the ArchivistaBox has been shut down (after the installation).

44.6 My CD does not boot

After downloading the update you must create a CD from which you can boot your computer. If your CD does not boot or if the message **No suitable CD found** appears, it may mean that your CD was not written as **image**.

Please write the ISO file as such and not as data CD that will never boot!

44.7 Installing ArchivistaBox on a Virtual Machine

ArchivistaBox is on the one hand a ready-made piece of hardware that you can buy from Archivista GmbH. On the other hand it is also a configuration that, in principle, turns any PC into an ArchivistaBox, i.e. an archiving server.

If you do not possess an unused PC waiting for a new life, there is the attractive alternative of an installation within a virtual box open to you. In the following it is described how to go about installing ArchivistaBox with the two products Vmware Player and VirtualBox.

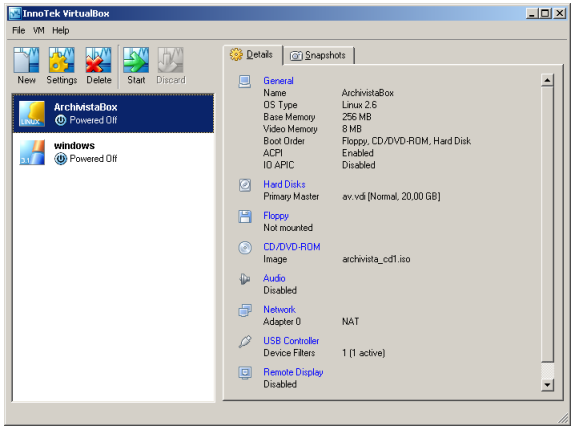
44.7.1 Installation with VirtualBox

Since January 2007 VirtualBox is available as OpenSource software. However, we cannot recommend anyone to compile VirtualBox him- or herself. Even if it worked out, the OpenSource version misses one crucial bit of software: the USB drivers. The pre-compiled version possesses the USB drivers and that is why we advise using the pre-compiled VirtualBox.

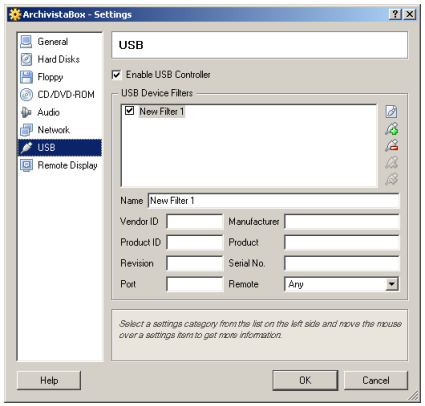
It has to be added that the pre-compiled version is restricted to personal use (whatever that means) and/or to testing purposes.

➲ Regarding the installation it may be said that the software can be downloaded from www.virtualbox.org (ca. 11MByte) and that version 1.3.4 (or higher) must be used as with previous versions (at least on our test computers) the USB ports did not run properly.

After the installation please start VirtualBox and set up a session according to the example below:



Please take special note of the USB port settings:



44.7.2 Installation with VMWare Player

Note: VMWare Player currently only offers USB 1.1 (situation as of February 2007). The use of a document scanner together with USB 1.1 can be recommended only with reservations, if at all.
Below a sketchy outline of the procedure is given.

- Download of VMWare Player from www.vmware.com/download/player/ (File size ca. 30MB).
- Installation of VMWare Player.

- Download VMWare image from www.archivista.ch (Zip file, file size ca. 10KB).

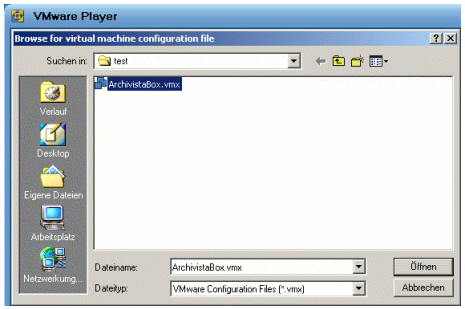
- Unzip this file into an empty directory. A hard disk with sufficient free space should be chosen: at least 5GB, 10GB would be better.

- The ArchivistaBox ISO file (ca. 600MB) that is also downloadable from www.archivista.ch must be copied into the same directory in which the VMWare image has its location.

- Start up VMWare Player by double clicking

.

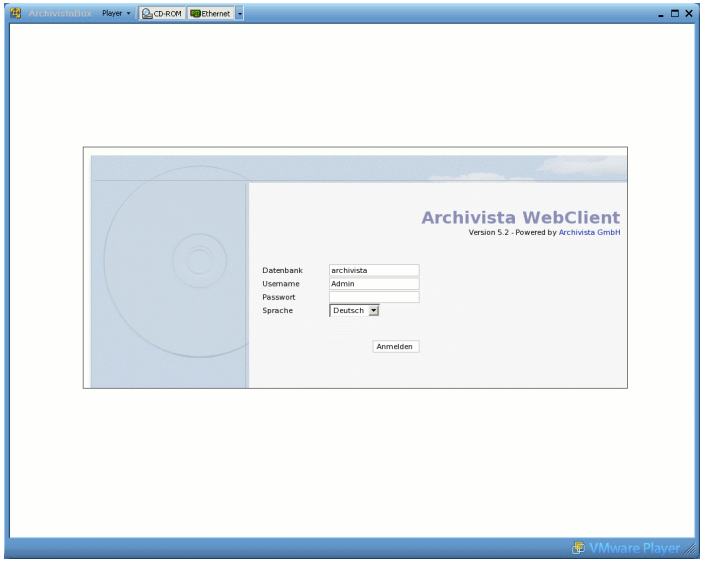
- Go to the said directory (where the unzipped file of the VMWare image is located and the ArchivistaBox ISO file). Open ArchivistaBox file.



- The ArchivistaBox starts up. Please wait a few minutes until you see the WebClient.

- It may happen that a message comes up saying that the RAM ID is not unique: 'If the virtual machine has been copied, you should create a unique identifier (UUID). If it has been moved, you should keep its old identifier.' Press 'create'.

This will make the ArchivistaBox start up. After a few minutes you can begin to try out the functionalities of ArchivistaBox.



44.8 Saving your data

➲ Please note the difference between long term and short term data saving. The ArchivistaBox concept supports the following levels of data security.

- Mirroring to second ArchivistaBox to save changes made during the current day (only ArchivistaBoxes 'Titlis' and 'Eiger').

- Daily backup by means of a) USB stick or b) tape

- Storing of data on non-rewritable media like CDs or DVDs for longterm data storage.

Mirroring to a second ArchivistaBox is described in 8.13. Saving your data by means of an USB stick is discussed in the next section.

For a description of the backup procedure by means of a tape we refer you to 8.8.1.

We recommend (and many countries require it) that you store your data on non-rewritable media for the purpose of safe **long-term keeping.** The procedure is described in more depth in 8.5.4.

44.9 Data exchange and backup by means of USB stick

44.9.1 General

We recommend that you buy USB sticks. USB sticks play an important role in the data exchange between ArchivistaBox and other digital devices and they are very easy to use.

Depending on the action you want to take you create a folder on the formatted stick (e.g. 'exchange') and you connect the stick with your ArchivistaBox. On the basis of the folder name ArchivistaBox will carry out the action you chose. In principle, an empty folder on a USB stick leads to an export from your ArchivistaBox while a folder containing data incites ArchivistaBox to import data. Currently, there are the following types of action:

pdf	PDF export of specific documents
exchange	Import/export of documents
crypt	Import/export (encrypted)
transfer	Import/export between box and RichClient
backup	Backup (for smaller archives)
cust	Cust folder (user specific scripts)
config	All parameters from an ArchivistaBox
'other name'	Import of images to Archivista

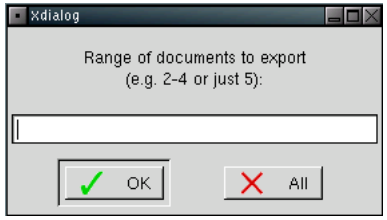
In the following the procedure is described for the various ways of exporting and saving data. The procedure for the saving of data as PDF files is described in detail. The rest works as described for PDF files.

44.9.2 PDF

As explained above the ArchivistaBox recognizes what action it is to take by the properties of the USB stick. For a PDF export to take place the USB stick must be prepared accordingly.
Go through the following steps:

- Attach the USB stick to your Windows PC
- It appears as 'Removable medium' on your PC
- Right click your mouse: format as Fat32
- Right click your mouse: create a directory / folder named 'pdf' on your USB stick (lower case)
- Right click your mouse: 'Eject'
- Remove the USB stick and insert it in the USB interface at the front of your ArchivistaBox

After entering the Archivista Admin user password you see the following dialog box:



Enter the range. (The figures 'from-to' refer to the document numbers.) Now the ArchivistaBox copies the documents onto the USB stick. When the process is finished the following message will come up:

44.9.3 Ordinary Archivista export for later re-import to ArchivistaBoxes

Proceed in the same way as described above in 'PDF export' but instead of creating a directory named 'pdf' create one called 'exchange'.

After you have entered the Archivista Admin user password the copying process is carried out. When the process is finished you will see a message. The export has succeeded.

44.9.4 Encrypted Archivista export for later re-import to ArchivistaBox

An encrypted export makes sense in those cases where you later want to import the encrypted data to another PC and decrypt them. In order to be able to make use of the encrypted export possibility a key must be available **before** the actual export process: you generate the key on the target PC and copy the public key to the source PC. To generate a key we refer you to 8.9.3.

Proceed in the same way as described above in 'PDF export' but instead of creating a directory named 'pdf' create one called 'crypt'.

After entering the Archivista Admin user password you see the following dialog box:

Now the ArchivistaBox copies the documents onto the USB stick. When the process is finished you will see a message. The export has succeeded.

44.9.5 Archivista export for later re-import to RichClient

Proceed in the same way as described above in 'PDF export' but instead of creating a directory named 'pdf' create one called 'transfer'.

After you have entered the Archivista Admin user password the copying process is carried out. When the process is finished you will see a message. The export has succeeded.

44.9.6 Backup

Proceed in the same way as described above in 'PDF export' but instead of creating a directory named 'pdf' create one called 'backup'.

After you have entered the Archivista Admin user password the copying process is carried out. When the process is finished you will see a message. The backup export has succeeded.

44.9.7 Customer specific adjustments (cust)

Each ArchivistaBox has the same structure. Sometimes an ArchivistaBox is supposed to fulfill a very specific function for a customer - a function which the customer does not wish to publish as OpenSource license or one which we cannot incorporate into the official ArchivistaBox (e.g. for reasons of space).

Further cases are: form recognition and specific programs for the processing of print data (COLD).

In all these cases we use the `cust` folder together with a USB stick. If the folder is empty, all available scripts/data are exported.

If the folder is not empty, the data from the USB stick are copied to the ArchivistaBox.

44.9.8 Image import from digital cameras (other folders)

If the stick does not show a specific folder, ArchivistaBox searches for JPG images and, if available, offers to import these. This function is described in detail in 45.8.

44.10 LCD display with ArchivistaBox

The LCD display includes the option of micro screen (4x20 characters) for the ArchivistaBox. This can be used for scanning and calling the system settings.

44.10.1 Scanning

In order to use the LCD display for scanning, connect it to a USB port on the ArchivistaBox. After a welcome message, the system displays the first scan definitions.

To trigger the scanning process, click on 'OK' on the current definition (labelled with a >). This triggers the scanning process immediately.

➲: Important: Scanning only works if you activate the scan button of the ArchivistaBox under 8.14.1.5.

To select a different scan definition, use the arrow keys located above and below the 'OK' button.

44.10.2 System settings

You can also use the LCD display to call the current system settings. To display the system settings, press the 'Home' key (located on the bottom right-hand side). The system settings are then displayed. You can use the arrow keys to scroll up or down.

➲: To return to the scan definitions from the system settings, press the 'ESC' key.

45 Integration questions

45.1 Is there a mail archiving function?

Yes. An IMAP-compatible mail archiving function that can monitor and archive any folders is available for the ArchivistaBox. Archived mail can also be checked out again. See 8.12.2 for more information.

45.2 Can I archive Office files (incl. check out)?

Yes. As of release 2009/I, the ArchivistaBox can process more than 200 file formats without using external plug-ins. See 8.11.1.1 or 12.1 for more information.

45.3 How to scan from one ArchivistaBox to the other

As we have seen in 8.11 one can scan directly to an ArchivistaBox by means of a network scan device. If there is no such device around, then ArchivistaBoxes lend themselves very well to this task of peripheral scan stations: from there the data are sent to a central server (e.g. ArchivistaBox EIGER).

In the following we outline the setup of such a scan station.

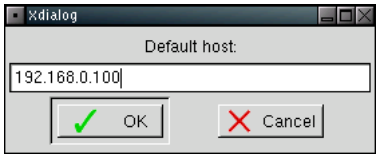
45.3.1 1st step: setting up the user on the server

If you have not done it already, you must specify the user who is to scan remotely into the database. Please make sure that you specify % for Host!

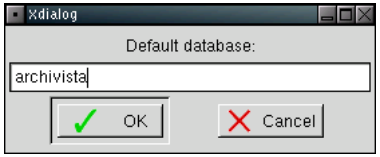
45.3.2 2nd step: setting up the scan station

It goes without saying that the scan station must be connected to the network. By means of the function 'Configure login mask' in menu 'System' you can now adjust the WebClient on the ArchivistaBox in a way that there is only **one** location to which one can scan, namely the database thus specified.

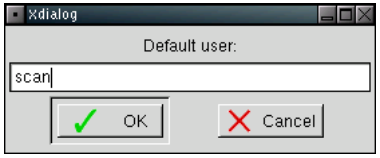
Proceed in principle as described in 8.14.1.4. In particular you must define that scanning is exclusively possible to a 'default host' and a 'default database'.



Under 'default host' enter the IP address of the server on which the database is.



Here you enter the database to which you want to scan from this scan station.



Lastly, you can specify the user in order that the main user need not enter her/his name each time she/he logs in. However, this name may be overwritten in the login mask.

We have now set up an ArchivistaBox as an ideal scanning station.

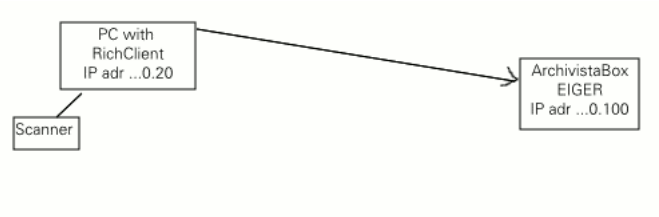
45.4 Scanning to ArchivistaBox with RichClient

It is possible to access an ArchivistaBox from a PC on which the RichClient is installed. Advantage: The scanning flexibility of the RichClient may be exploited. For example, metadata of an existing document can be copied and re-used for the next document (something which is not possible in the WebClient).

There are three alternatives, which are described below.

➲ Please take note of the IP addresses in the illustrations. For the sake of clarification they are repeated in the screen shots.

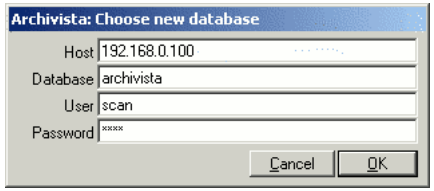
45.4.1 Scanning by means of RichClient and Windows scanner



The first alternative concerns a typical installation under Windows with a scanner run by a TWAIN driver and attached to a PC. Scanning takes place locally and the images are sent to the server (here: ArchivistaBox EIGER).

45.4.1.1 Logging in

The IP address of the ArchivistaBox to which the scans are ultimately sent must be entered. Also the database name must be entered. It must be borne in mind that the user has to be one that is allowed access not only locally. See also 35.2.1.2.



45.4.1.2 Displaying the scanning form and adding pages

In menu 'Page' in Page view mode there is the command 'Scanning with settings... (F2)'. With this command pages can be added. They are stored on the server (here: ArchivistaBox EIGER).

➲ Please note that the option 'Save added pages directly in database (instead of input folder)' must be activated (Main view, menu 'Database', Parameters / Archiving & index...).

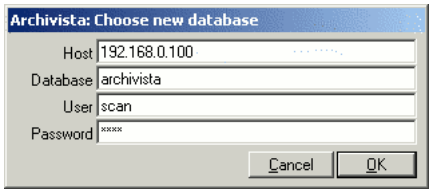
45.4.2 Using the RichClient to scan to server to which scanner is attached

The illustration for the second alternative shows that the scanner is plugged into the server (here: ArchivistaBox EIGER). The scanning process is started from a separate PC on which the RichClient is installed.

45.4.2.1 Logging in

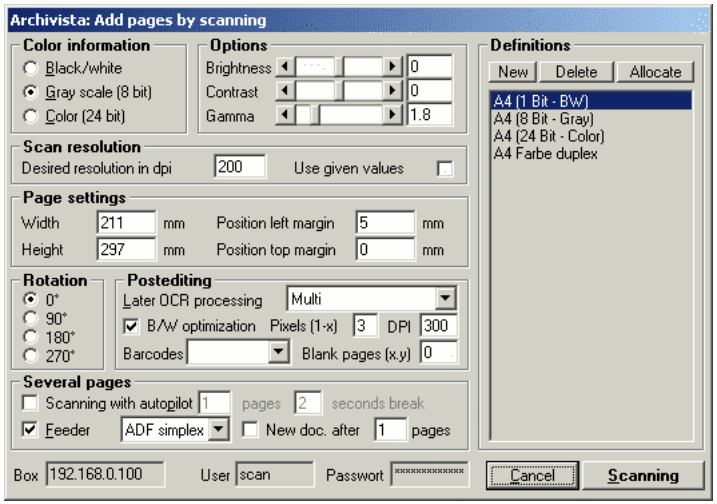
The IP address of the ArchivistaBox to which the scans are ultimately sent must be entered. Also the database name must be

entered. It must be borne in mind that the user has to be one that is allowed access not only locally. See also 35.2.1.2.



45.4.2.2 Displaying the scanning form and adding pages

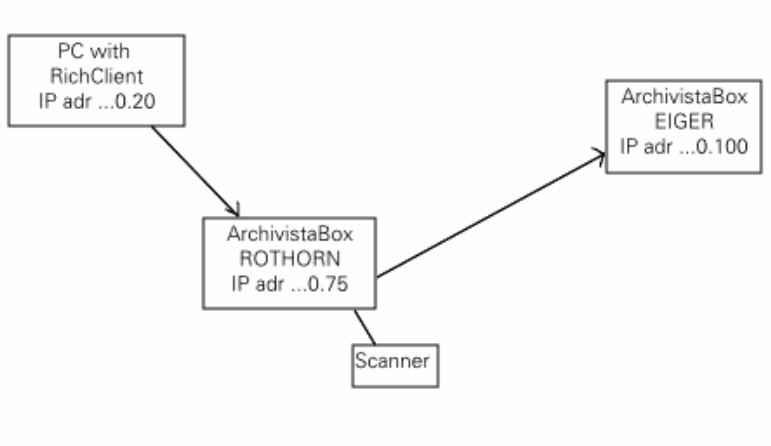
In menu 'Page' in Page view mode there is the command 'Scanning with ArchivistaBox... (Shift+F1)'.



At the very bottom the IP address of the ArchivistaBox that is used as scanning station is entered (here ArchivistaBox EIGER). The scan definitions that are displayed are those specified for the database 'archivista' on the server (here: ArchivistaBox EIGER).

➲ Please note that the option 'Save added pages directly in database (instead of input folder)' must be activated (Main view, menu 'Database', Parameters / Archiving & index...).

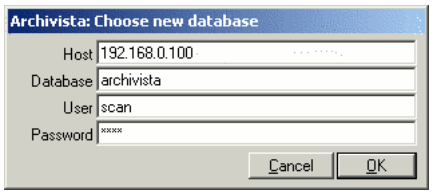
45.4.3 Using the RichClient to scan to server via scanning station



As the illustration for this alternative shows the scanning process is triggered from a PC with the Archivista 2009/I RichClient but the scanning station itself is a ArchivistaBox with attached scanner (here: ScanBox ROTHORN). The scanned pages are stored on the server in the end (here: ArchivistaBox EIGER).

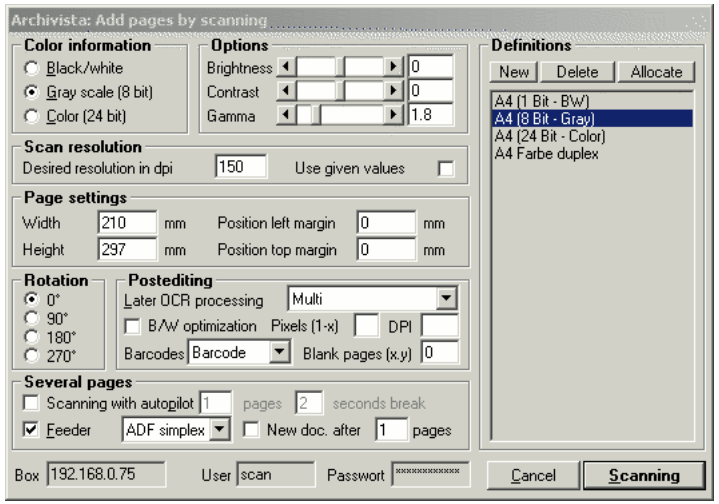
45.4.3.1 Logging in

The IP address of the ArchivistaBox to which the scans are ultimately sent must be entered. Also the database name must be entered. It must be borne in mind that the user has to be one that is allowed access not only locally. See also 35.2.1.2.



45.4.3.2 Displaying the scanning form and adding pages

In menu 'Page' in Page view mode there is the command 'Scanning with ArchivistaBox... (Shift+F1)'. When you run it, a special form or dialog box appears.



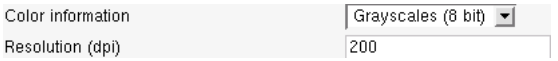
At the very bottom the IP address of the ArchivistaBox that is used as scanning station is entered (here: ArchivistaBox ROTHORN). In addition, a specific user must be named that was previously created on the ScanBox with IP address ...0.75. (Please refer to 8.14.1.7 regarding the proper procedure). The scan definitions that are displayed are those specified for the database 'archivista' on the server (here: ArchivistaBox EIGER).

45.5 Can I scan books?

A book can be scanned by means of the usual scanning command. Special attention should be paid, however, to a number of points. To create a scan definition for a book go to WebAdmin and menu **Scanning**. Click **New** to make a scan definition that fits the specific needs of the book to be scanned. For a detailed description of all scan definitions please read Chapter 18.

When scanning books it is fundamental to give the exact format. The scale of measurement is in millimeters. Inexact specifications may lead to a considerable impairment of scanning quality. Enter the measures in the following fields:

As most books are printed in black and white, the specifications **Grayscales** and 200 dpi are sufficient. If the books are printed in color, these values must be adjusted accordingly.

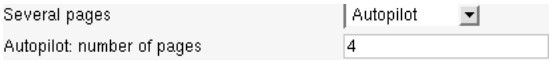


To achieve a really high scanning quality the option **Black/white optimization** should be ticked.

Maybe you are asking yourself why you do not scan in black and white. The answer is that black and white optimization does not work with a black and white scan.

Depending on how you scan your book it must be rotated in order that the scan is in an upright position later. For this purpose use the field **Rotation**. The scan is turned automatically in accordance with your settings.

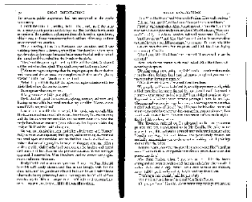
The option **'Several pages'** relieves you from the necessity to start the scanning command again and again. Under **Autopilot** one can specify how many scans should be made.



In the example above a '4' has been entered in the field **Autopilot: number of pages**. The scanner makes 4 scans. Between the scans the scanner takes breaks. By means of **Autopilot: break in seconds** it can be specified how long that break ought to be.

The illustration below shows an example where the scanner is to wait 2 seconds in between scans.

When scanning a book you mostly scan two pages at once.



The function **Split page in middle** gives you the possibility to split the pages in two after scanning.

Since books are in most cases made up of text one should not forget **text recognition (OCR)**. Under 'Desired OCR definition' you can choose between Multi, German, English and German, and others.

45.6 Classifying documents (ERP interface)

You use the Archivista printer to print documents in the archive. You then want to index these documents automatically and you use Archivista to search for a solution. This solution is called 'coldplus.pl'.

`/home/data/archivista/cust/cold/coldplus.pl`

If this program exists in the ArchivistaBox, it is run for each print job. For example:

```
#!/usr/bin/perl

my $db = shift; # database name (you can change it)
my $title = shift; # title of document (goes to Title)
my $pdffile = shift; # pdf file name (for own work)
my $psfile = shift; # ps file name (for own work)
my $for = shift; # author (goes to Owner)

if (length($title)>120) {
  $title = substring($title,0,119);
}
$title =~ s/Microsoft Word - //g;
$title =~ s/Microsoft Office //g;
$title =~ s/;/ /g;
$title =~ s/:/ /g;
$title =~ s/=/ /g;

# to return values, you need to print it
# at first position include db name, followed by a ;
# after it add the field=value, separated with a :
my $ret = "$db;Title=$title:owner=$for";
print $ret;
```

The program always receives the database name, the respective PDF file, the corresponding postscript file and the author (if the author is included in the PDF document). The program initially returns the database name as well as the desired fields.

➲: ftpplus.pl exists in addition to coldplus.pl. The difference here is that only the name of the document is received. You have to add the corresponding program to the following location in the ArchivistaBox:

`/home/data/archivista/cust/ftp/ftpplus.pl`

45.7 Is there an API for the ArchivistaBox?

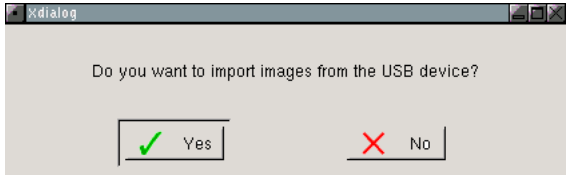
Yes. Use the external control for the Archivista WebClient. See 39 for more information.

45.8 Import of images

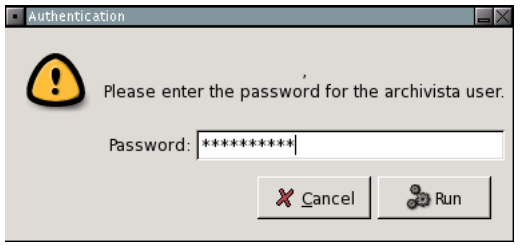
There is the possibility to store images from a digital camera directly on your ArchivistaBox. Attach your digital card or a USB flash SD card reader to the USB interface on the front of your ArchivistaBox.

➲ In general we recommend using a USB flash SD card reader for data transfer.

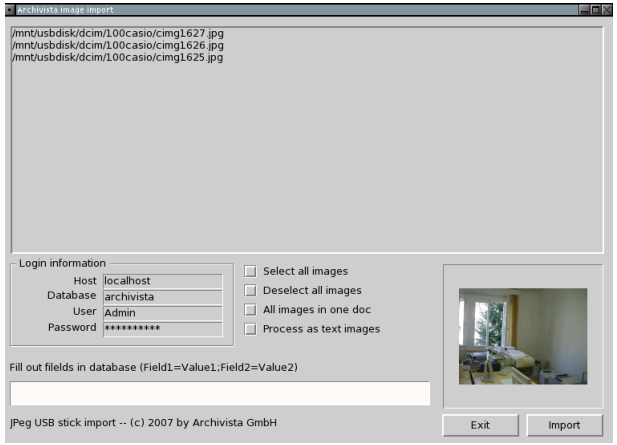
This kind of data transfer gives you higher convenience and flexibility. In addition, the direct transfer from camera to ArchivistaBox may require special settings which the ordinary user may not be able to see through without disproportionate effort. When the computer notices that you want to import image files you will see the following screen:



Click **Yes** to import the images. Then enter your password:



Now you get to the main menu:



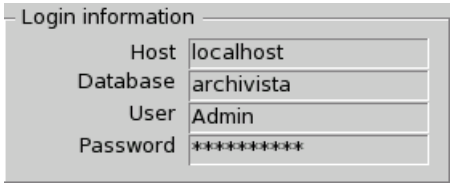
You have the following choice:

- **Select all images**
- **Deselect all images** (= undo selection of all images)
- **all images in one doc** (= file all images in one document)
- **process as text images** (= import photographed text pages)

The field **Fill out fields in database** serves for the allocation of keywords to the individual photos. You may enter something like: Title=Office.

Subsequently, the column 'Title' in the WebClient will show the name of the image: **Office**.

Enter the name of the column in which you want to see the keyword. Enter the equals sign and then add the keyword you want. In the bottom left corner you see the following dialog box:



Normally this box shows the address of your ArchivistaBox. If you want to access a different ArchivistaBox you can change the data here.

To import the images click **Import** and, after that, **Exit** to leave the menu. A message appears telling you that you can remove the USB medium. Click **OK**.

The WebClient gives you the possibility to change to the so-called picture mode. Click the **tree symbol** in the top bar.

Your photos will appear as in a photo gallery, the evident advantage being that you can view all photos at once and can easily switch from one to the other. You can leave the picture mode by again clicking the **tree symbol**.

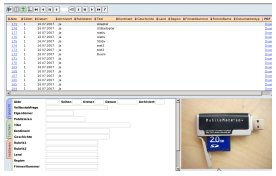
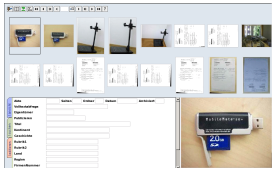
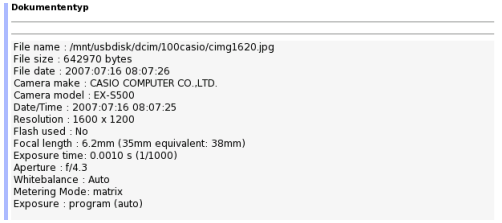


TABLE MODE


PICTURE MODE

What is more, the WebClient - both in picture mode and normal mode - lets you retrieve the **Exchangeable Image File Format Information**, i.e. the EXIF information. Simply click the blue **View** tab.

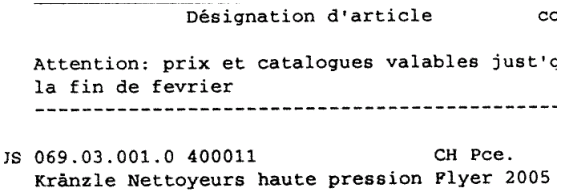

Dokumententyp
File name : /mnt/usbdisk/dcim/100casio/cimg1620.jpg
File size : 642970 bytes
File date : 2007:07:16 08:07:26
Camera make : CASIO COMPUTER CO.,LTD.
Camera model : EX-S500
Date/Time : 2007:07:16 08:07:25
Resolution : 1600 x 1200
Flash used : No
Focal length : 6.2mm (35mm equivalent: 38mm)
Exposure time: 0.0010 s (1/1000)
Aperture : f/4.3
Whitebalance : Auto
Metering Mode: matrix
Exposure : program (auto)

45.8.1 Photos with text

The import of photographed text documents works in analogy to the import of photos, however, there is one difference. In the main menu of the photo import feature click **Process as text images**! The photo is imported as usual and is archived by ArchivistaBox. Prerequisite for a good 'scan' of the text document is a good quality, high-resolution camera. Many digital cameras possess a **text mode** specifically intended for the photography of texts: go to the menu 'settings' of your camera. Additionally, a camera support may be a definite advantage as pictures taken without support are easily blurred.

Such a camera support may be purchased from Archivista GmbH.

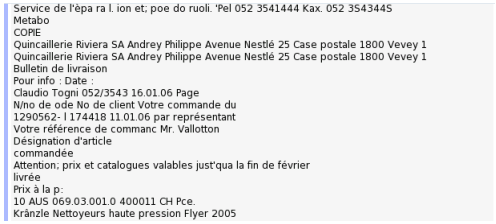
A camera of good quality (with, say, 10 megapixels, but even 5 megapixels may be sufficient) will return a photographed document of excellent quality. Let yourself be convinced by the following example:


Désignation d'article cc

Attention: prix et catalogues valables just'c
la fin de fevrier

JS 069.03.001.0 400011 CH Pce.
Kränzle Nettoyeurs haute pression Flyer 2005

If the quality is sufficient (the letters must be closed and easily readable), the text recognition works also with a photographed document and the text of the document can be found behind the blue **View** tab in the WebClient.


Service de l'épa ra l. ion et; poe do ruoli. 'Pel 052 3541444 Kax. 052 3S4344S
Metabo
COPIE
Quincaillerie Riviera SA Andrey Philippe Avenue Nestlé 25 Case postale 1800 Vevey 1
Quincaillerie Riviera SA Andrey Philippe Avenue Nestlé 25 Case postale 1800 Vevey 1
Bulletin de livraison
Pour info : Date :
Claudio Togni 052/3543 16.01.06 Page
N/no de ode No de client Votre commande du
1290562- I 174418 11.01.06 par représentant
Votre référence de commanc Mr. Vallotton
Désignation d'article
commandée
Attention; prix et catalogues valables just'qua la fin de février
livrée
Prix à la p:
10 AUS 069.03.001.0 400011 CH Pce.
Kränzle Nettoyeurs haute pression Flyer 2005

45.9 Transferring the RichClient to the ArchivistaBox

The easiest way to transfer an existing RichClient database to the ArchivistaBox is to carry out an export from the old Archivista database and then an import using a USB stick. You have to save the exported data in the 'transfer' folder. This is the only way to ensure that the USB stick is recognised correctly in the ArchivistaBox.

➲: Please note that the field information is only transferred if the database in the ArchivistaBox has the same fields as those of the old RichClient database.

45.10 Importing data from other DMS software

The ArchivistaBox has **fully automated import functions.** In order to carry out a transfer, the data has to be exported from the old system. You can then create an import file and then import it to the ArchivistaBox either by using ArchivistaGPL or the USB stick ('transfer' folder). For information about the import format, see 30.3.1.

➲: You can easily identify the format to be used by exporting two or three files from the ArchvistaBox or ArchivistaGPL.

45.11 ArchivistaGPL (RichClient)

45.11.1 Under which operating system does Archivista 2009/I run?

Archivista runs under 'Windows'. With Archivista 2009/I you can use Windows9x/Me and WindowsNT4/2000/XP or higher. Provided you work with Wine on Linux you can run Archivista on Linux, too.

A note for specialists: Archivista 2009/I was developed with VisualBasic Professional. As Database server Archivista 2009/I uses MySQL 4.0.x. Large parts of the additional modules (e.g. the WebClient) were realised in Perl; these programs can be run both on Windows and on Linux without any restrictions.

45.11.2 What are the system requirements for the installation of ArchivistaGPL?

In order to work efficiently with Archivista 2009/I you need a Pentium PC with a minimum of 64 MByte RAM. Further you need Windows9x/Me or WindowsNT4/2000/XP as operating system.

A scanner is not imperative (you can create pages directly from graphic files), nevertheless the use of a scanner makes good sense and sooner or later you will need one. The same applies to a good printer and a CD writing device. The latter is used to transfer the high resolution pages (which are needed for top quality printouts) onto CDs.

45.11.3 What do I have to take into consideration when converting from an earlier version?

Converting Archivista 4.x/5.x: Please see the information under 45.9.

45.11.4 Are there any fees if I migrate to version 2009/I?

Yes! Archivista has been on the market for more than 5 years. Previous updates have been free. Since Version 2009/I is a quantum leap in the history of the product we ask for a migration fee. The exact prices can be found at www.archivista.ch. Alternatively, you can run ArchivistaBox under the General Public License (GPL).

45.11.5 Which archiving formats does Archivista 2009/I offer and how do they differ?

Archivista 2009/I offers four formats. In the following we give an overview by comparing the four.

Encrypted (Bmp/Zip): the standard format consists of compressed Windows bitmap files. **Target group note:** Don't use it any longer.

Bit Black/White (Tiff G4): Widely used file format in archiving systems. **Target group note:** Open archive in black and white without any need for confidentiality.

Loseless (Png): Archivista standard format with excellent compression rate with black/white, gray scales and color pages. No quality loss! **Target group note:** Don't use it any longer.

Compressed (JPeg): Excellent compression rate (e.g. 1:20), however, not without any losses. Depending on the compression factor you must live with a considerable quality loss. **Target group note:** Archiving of color photographs.

➲ The archiving format may be specified for each individual document in field 'Format' in the main view. However, you must make your choice of format **before** adding pages to the document.

45.11.6 Why is it impossible to add pages after a document has been archived?

During the archiving process the documents are prepared for transferral to external storage devices. The pages are compressed and moved to a folder in the Output directory. I.e. after the archiving process is finished and after a folder is closed the documents may be transferred to an external storage device such as a CDR (see also chapter 8.5.4).

If it were possible to add pages later, a folder could never be closed and consequently could not be transferred. That is the reason why documents may not be extended later.

➲ To circumvent this problem work with document templates. First search for the document to which you would like to add pages. Create a document template by copying (Ctrl+C) and pasting the document (Ctrl+V). You will get a second document which corresponds to the first: the most important fields show the same information. Scan your pages into this new document. Later when selecting your document by keyword, title or persons both documents will be shown. There is the considerable advantage that you may sort your documents according to date.

45.11.7 Can I customize the look of the start-up screen?

Yes, you can set preferences with regard to the look of the start-up screen and the printing options. The **size of the program window** is to be found in file **'AV41.SCR'.** If you want the program window to fill the entire screen at start-up, change the value in line one from '-1' to '0'.

Further options regarding the display of **pages and printing options** can be found in file **'AV41.STA'.** You will find many settings influencing the view of the start-up window there. Take note of line two and the word **'Activated'.** In this context '0' means that user defined settings are ignored, '1' means that the settings in this very file determine the looks of Archivista at start-up and the default value '2' means that the settings are kept from one session to the next.

45.11.8 Can I export my data?

Yes, by using the function 'Export data' in chapter 30.3.1. We would like to point out that you can copy any marked fields by pressing **'Ctrl+Insert'** and paste them in any application (also a text application) with **'Shift+Insert'.** This also works with RTF fields.

45.11.9 Can I transfer archived documents from one Archivista database to another?

Use the export and import functions for this process by including the pages. Do make sure that all (uncompressed) pages find room on the relevant hard disk and in the relevant directory.

45.11.10 Printing problems

Should you get funny page breaks after printing with Archivista, the problem lies probably with the minimal margins of the printer. Please adjust the relevant values: either in the general parameters or in the print reports (Menu 'Database' and 'Parameters').

With some printers printing of the last page takes longer. You can avoid this by starting another print job.

45.11.11 Is Archivista web-enabled?

Yes! The simplest and most efficient method to put Archivista databases on the World Wide Web is the ArchivistaBox. See Part III.

45.11.12 Which scanners work with ArchivistaGPL?

You should use an ArchivistaBox for scanning. Since 1.1.2008, Archivista has only provided scanner support for the ArchivistaBox. You can find a list of the scanners that are currently supported on our homepage www.archivista.ch.

If you want to use a scanner in Windows with ArchivistaGPL anyway, it is still possible to address the scanner in ArchivistaGPL via a TWAIN interface. Unfortunately, however, many problems occurred with TWAIN scanners throughout the past year.

One final option is to move the copy button of the delivered scanner software to an activated Archivista printer. See 8.10.1 for information on how to set this up. If you want to print scanned documents in bleed-off form, you should use a printer driver in Windows that is capable of handling this.

46 Solutions for common problems

46.1 Scanning is slow

Scanning may be slowed down by some of the options. Functions that use considerable calculating capacity include:

- OCR recognition
- PDF generation
- Barcode recognition (optional module)
- Removal of blank pages

If several of these are turned on at the same time and the system is supposed to go through all tasks at once scanning speed may sink. If you want all of these functions, turn them off during scanning and turn them back on afterwards.

46.2 Scanner is not working

The first question you should ask yourself is: am I using the right scan definition? If your papers are sitting in the ADF but you are using a scan definition where 'ADF' as option is not chosen, then the scanner cannot respond.
The second possibility may be that the connection between ArchivistaBox and scanner is broken.
The scanner AVision AV220 with USB interface should immediately respond when a connection with the mains on the one hand and the ArchivistaBox on the other has been established. A startup procedure where the scanner has to be turned on first and the ArchivistaBox afterwards is not required. However, sometimes this connection is interrupted without any apparent reason. In such a case simply turn the scanner off, and after a minute turn it back on. In 9 out of 10 cases the connection will be resumed. If not, try as a last resort shutting down the ArchivistaBox and starting it up again. See 44.3 und 7.5.
The scanner AVision AV3200c is connected to the ArchivistaBox by means of a SCSI card. If scanning by means of this device is no longer working, please refer to the function described in 8.14.1.6.

46.3 Characters on keyboard do not work

If you cannot type characters that are represented on the keys of your keyboard, then you must adjust the keyboard layout setting. See above in 8.14.1.1.

46.4 OCR is not (or no longer) functioning

Should the text recognition not be working, the underlying causes may be the following:

- OCR module has not been registered
- OCR has not been activated for this particular database
- A 'warped' log entry prevents the OCR software from working

The measures below should put things right (order according to the causes listed above):

- OCR module must be registered; see 8.14.4.5
- OCR must be turned on for the database in question; see 17.1
- Delete log table and restart OCR; see 8.14.4.2 and 8.14.4.4

46.5 ArchivistaBox is no longer working

ArchivistaBoxes are designed for 7x24 performance and we recommend that you keep them running - over months and years.

If an ArchivistaBox suddenly fails and the problem is not due to the software, you can try to solve the problem by shutting down the mini PC and starting it up again. This means that this is one of the rare cases when it makes sense to actually touch the hardware.

Follow the procedure as described in 44.3 above.

If this does not work, then press the 'Power' button for ca. 7 seconds. It is to be found on the front of the Archivista box. Press it until the appliance turns itself off and the little lamps go out.

After the ArchivistaBox has rested for a few minutes press the 'Power' button again. The system should start up now. If it does not, please contact Archivista GmbH.

46.6 Box is running but the monitor is black

Try switching the monitor off and back on.

46.7 Box is running but programs have vanished

If all programs have been closed down and a grey screen is all that you see, press 'F10' to reset the display.

46.8 A user has forgotten their password

You can reset a user?s password via WebConfig. See the information under 26.8.

46.9 'Document ... locked' message is displayed

If you do not exit the program properly (i.e. by pressing Ctrl+Alt+Del to close the application), a document that is undergoing processing may be locked. You can unlock documents that are locked in this way via WebConfig under 'Unlock documents'. See the information under 26.7.

Index

Abbreviations, 254
Abbreviations - RichClient
    working with, 163
Access rights - RichAdmin, 197
Adding pages - RichClient, 153
Additional fields - RichAdmin, 211
Additional fields - RichAdmin, 210
Address of Archivista GmbH, 9
Addresses - RichClient, 189
Alt+F4 - RichClient, 162
API
    Applications, 249
    Archivista tables, 245
    Login to databases, 249
Applications
    own ones, 249
Archive - Introduction
    folder, 20
    pages, 20
Archive administration - WebAdmin, 103
Archive search mode (Ctrl+F9) - RichClient, 170
Archive, digital - Introduction, 18
Archiving - RichAdmin, 203
Archiving concept, 245
Archiving process - ArchivistaBox, 43
Archiving & index - RichAdmin, 202
Archivista document, the - Introduction, 20
Archivista-BarcodePrint - RichClient, 114
ArchivistaBox, 34
Auditing acceptability - ArchivistaBox, 261
Auxiliary tables - RichClient
    Rubric (Literature), 191

Backup - ArchivistaBox, 49
Barcode - RichClient
    Code 39, 114
    Interleaved 2 of 5, 114
Barcode processing - RichClient, 113
    Why?, 113
Barcode technology - RichClient, 113
Barcodes - WebAdmin, 113
Batch paths for network document servers - RichAdmin, 202
Batch paths for Xerox Work Centres - RichAdmin, 202
Black rectangle - RichClient, 156
Blanking fields, 211
Blanking fields - WebAdmin, 101

Client - Introduction, 18
Clipboard - RichClient, 163
Computer files - Publishing Edition, 231
Controlled vocabulary - WebClient, 82

Conversion, 272
Copy commands under Windows, 254
Copyright, 13
Creating a document - RichClient, 152
Creating and scanning documents - WebClient, 86

Database - Introduction, 18
Database - RichClient, 158
    choosing other, 159
Database conversion, 272
Database creation - WebAdmin, 136
Defining fields & masks - RichAdmin, 209
detail view, table 'Archive', 154
Digital archive - Introduction, 18
Directories - RichAdmin, 201
Document - RichClient
    Creation and scanning, 167
Documents, 20
Documents - RichClient
    create, copy, paste and delete, 164
Download - WebClient, 87

E-mails, archiving of - ArchivistaBox, 59
Edit - RichClient, 163
Edit mode - WebClient, 82
Editing and deleting pages - WebClient, 87
Efficient entering of data, 254
Electronic documents, archiving of - ArchivistaBox, 54
Electronic files - Publishing Edition, 231
Encryption - ArchivistaBox, 51
Entry conditions 'archive' - RichAdmin, 193
Entry conditions - RichAdmin, 218
Error 3024, 255
Error messages, 255
Exporting data - RichClient, 159
Exporting pages - RichClient, 181

FAQ - ArchivistaBox, 257
FatClient, 152
Field definition - WebAdmin, 97
Field definitions and input masks - RichAdmin, 209
Field names - RichAdmin, 204
Field objects - RichAdmin, 212
Field type: 'Linked field', 83
Field types - WebClient, 82
Fields - Introduction, 19
Fields, special - RichClient, 157
File upload, 57
Filter - RichClient, 172
Folders, 248
Folders - Introduction, 19
Format - RichClient, 168
FTP-Server - ArchivistaBox, 56
Full text indexing with MySQL - RichAdmin, 204
full text search, 9
Function keys - ArchivistaBox, 73
Functions - RichClient, 156

Handbuch - WebConfig, 138
Help, 9
Help - RichClient, 174
Hyperlinks - RichClient, 160

Icons - WebClient, 78
Importing data - RichClient, 160
Initial start-up - ArchivistaBox, 37

Input masks - RichAdmin, 219
Installation - Introduction, 22
Internal image copies - RichAdmin, 204
Internet web page, 9
Introduction, 9

Key in red color - RichClient, 161
Keywording, 254

Licenses - RichAdmin, 195
Links between records - RichClient, 160
list, 154
Literature - RichClient, 190
Log in With New Password - RichClient, 158
Login - WebClient, 76
Login and logout - WebAdmin, 90
Long-term considerations on documents - Publishing Edition, 226

Macros - RichAdmin, 207
Macros - RichClient, 173
Macros- RichClient, 188
Mail archiving - WebAdmin, 134
Mail server - ArchivistaBox, 59
Main view and page view - WebClient, 76
Main view of table archive - RichClient, 154
Manual - Introduction, 9
Manuals - ArchivistaBox, 42
Mask definition - WebAdmin, 98
Master-Slave configuration - ArchivistaBox, 60
Memo fields - RichClient, 168
Menu - RichClient, 156
    Database, 158
    Edit, 163
    Macros, 173
    Search, 169
    Table, 173
Menus and toolbar - RichClient, 154
Modules
    Creating your own, 254
Modules - ArchivistaBox, 46
Mouse
    right key - RichClient, 157
Mouse - RichClient
    doubleclick page image, 175

Navigation - RichClient, 164
Notes - RichClient, 190

OCR - RichClient, 161
OCR definitions - WebAdmin, 123
Office files, 57
online help, 9
Online update, 70
online-tutorials, 152
Owner - RichAdmin, 198
Owners - RichAdmin, 199

Page settings - RichAdmin, 220
Page view - RichClient, 175
    exporting pages, 181
    Importing pages, 181
    menu 'Archive', 175
    menu 'Viewing', 186
    menu Document, 177
    menu Page, 178
    scanning, 178

Page view- RichClient
    macros, 188
Pages, 20
Parameters - RichAdmin, 193
    General program parameters, 193
Passwords - ArchivistaBox, 68
Passwords - RichAdmin, 197
PDF files - Publishing Edition, 231
Picture mode - RichClient, 168
Preliminary note - Publishing Edition, 228
Print reports - RichAdmin, 220
Print reports - RichClient, 158
Printing (Ctrl+P) - RichClient, 158
Printing - RichAdmin, 220
Printing - RichClient, 175
Printing in page view - RichAdmin, 194
Printing pages - WebClient, 88
Publishing - WebClient, 84
PublishingEdition, 226

Quit program - RichClient, 162
Quitting - RichClient, 162

Record - Introduction, 19
Record - RichClient, 164
Remote access - ArchivistaBox, 73
Replace - RichClient, 166
Report module - RichAdmin, 220
Resizing the detail view window (table 'Archive') - RichClient, 156
RichClient, 152
Right mouse key - RichClient, 157
RTF fields - RichClient, 168
Runtime errors, 255

Saving your data - ArchivistaBox, 261
Scan definitions - WebAdmin, 108
Scanning (F2) - RichClient, 178
Scanning with network scanner - ArchivistaBox, 56
Search (Shift+F5) - RichClient, 165
Searching - RichClient
    Documents (fields), 169
Searching - WebClient, 80
Selection - RichAdmin, 205
Selection - RichClient
    all records (F6), 172
    working with defined criteria, 172
Self-made applications, 249
Server - Introduction, 18
Settings - ArchivistaBox, 42
Shortcuts - RichClient, 156
Show pages - RichClient, 167
Shut down - ArchivistaBox, 73
Sort - RichClient
    ascending (Ctrl+F6), 172
    descending (Ctrl+F7), 172
Source formats - Publishing Edition, 227
SQL definitions - WebAdmin, 125
Starting up the first time - Introduction, 24
Starting up the program - RichClient, 152
status bar, 154
Storage concept - RichAdmin, 203
Store page (Ctrl+S) - RichClient, 180
Switching it on - ArchivistaBox, 38
Sysop privileges - RichAdmin, 193
System - ArchivistaBox, 64

Tab key - RichClient, 156

Table - RichClient
    Literature, 190
Tables - Introduction, 19
Tables - RichClient
    Addresses, 189
    notes, 190
Tiff-Multipage - RichClient, 181
Tips for archiving - Introduction, 20
Toolbar - RichClient, 156
Transferring a database to the ArchivistaBox, 272
Tutorial, 226

Update, 70
User administration - RichAdmin, 195
User administration - WebAdmin, 91
User manual - WebClient, 76

Versions of Archivista, 10
View - WebClient, 80

Web pages - Publishing Edition, 232
WebAdmin, 90
WebClient - ArchivistaBox, 42
Working method - Introduction, 19